BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 1993 by The American Society for Biochemistry and Molecular Biology, Inc.

Vol. 268, No. 21, Issue of July 25, pp. 15983-15993, 1993
Printed in U.S.A.

Erythropoietin Structure-Function Relationships

MUTANT PROTEINS THAT TEST A MODEL OF TERTIARY STRUCTURE*

(Received for publication, February 12, 1993, and in revised form, April 2, 1993)

Jean-Paul Boissel, Woan-Ruoh Lee, Scott R. Presnell‡, Fred E. Cohen‡, and H. Franklin Bunn

From the Hematology/Oncology Division, Brigham and Women's Hospital, Harvard Medical School, Boston Massachusetts 02115 and the ‡Departments of Medicine, Pharmaceutical Chemistry, Biochemistry, and Biophysics, University of California, San Francisco, California 94143-0446

On the basis of its primary sequence and the location of its disulfide bonds, we propose a structural model of the erythropoietic hormone erythropoietin (Epo) which predicts a four α-helical bundle motif, in common with other cytokines. In order to test this model, site-directed mutants were prepared by high level transient expression in Cos7 cells and analyzed by a radioimmuno assay and by bioassays utilizing mouse and human Epo-dependent cell lines. Deletions of 5 to 8 residues within predicted α-helices resulted in the failure of export of the mutant protein from the cell. In contrast, deletions at the NH_2 terminus (Δ2–5), the COOH terminus (Δ163–166), or in predicted interhelical loops (AB: Δ32–36, Δ53–57; BC: Δ78–82; CD: Δ111–119) resulted in the export of immunologically detectable Epo muteins that were biologically active. The mutein Δ48–52 could be readily detected by radioimmunoassay but had markedly decreased biological activity. However, replacement of each of these deleted residues by serine resulted in Epo muteins with full biological activity. Replacement of Cys^{29} and Cys^{33} by tyrosine residues also resulted in the export of fully active Epo. Therefore, this small disulfide loop is not critical to Epo's stability or function. The properties of the muteins that we tested are consistent with our proposed model of tertiary structure.

Humoral regulation of red blood cell production was first proposed at the beginning of this century (1). Convincing physiologic experiments documenting the existence of erythropoietin (Epo)[1] (2–5) were followed by its purification (6) and partial structural characterization (7). The molecular cloning of this biologically and clinically important cytokine (8, 9) has led to further understanding of its properties (10, 11).

The binding of Epo to its cognate receptor (12) on erythroid progenitors in the bone marrow results in salvaging these cells from apoptosis (13), allowing them to proliferate and differentiate into circulating erythrocytes. The Epo receptor is a member of an ever enlarging family of cytokine receptors (14). In like manner, Epo shares weak sequence homology with other members of a family of cytokines which also include growth hormone, prolactin, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, G-CSF, GM-CSF, M-CSF, oncostatin M, leukemia inhibitory factor, and ciliary neurotrophic factor (15–17). The genes encoding these proteins have similar numbers of exons as well as a clear relationship between intron-exon boundaries and predicted α-helical structure. These similarities have led to the prediction that this family of cytokines share a common pattern of folding into a compact globular structure consisting of four amphipathic α-helical bundles. Such theoretical models of the structures of human growth hormone (18) and IL-4 (19) have been in remarkably good agreement with subsequent structures established by x-ray diffraction (human growth hormone) (20, 21) or by multidimensional NMR (IL-4) (22, 23). Moreover, the crystal structures of GM-CSF (24) and monomeric M-CSF (25) are also in reasonable agreement with their predicted structures.

Thus far, the structure of Epo has not been analyzed by either x-ray diffraction or by NMR. In order to begin to gain an understanding of structure-function relationships, we have taken a three-pronged approach.

(*a*) Sequence determination of Epo from mammals of different orders in order to establish regions of homology (26).

(*b*) Construction of a model of the three-dimensional structure of Epo, followed by the design and preparation of muteins that test this model. These experiments are presented in this paper.

(*c*) Design and testing of muteins that provide information on receptor binding domain(s). This work will be presented in a subsequent paper.

MATERIALS AND METHODS

Computer-based Modeling of Structure

Prediction of Secondary Structure—Epo sequences from human, monkey, mouse, rat, sheep, pig, and cat were aligned (26) and examined using a hierarchical approach to secondary structure prediction that assumes that these proteins are members of the α/α folding class (27). First, the pattern-based method of Cohen *et al.* (28) for turn prediction was used to delimit sequence blocks likely to contain secondary structure. Predictions using the methods of Garnier *et al.* (29) and Chou and Fasman (30) suggested α-helical regions within these blocks. Finally, helical wheel projections were used to examine and then limit helix length based on preserving amphipathic character as codified in the work of Presnell *et al.* (31). The locations of glycosylation sites were also used to suggest helix boundaries.

Tertiary Structure Prediction—Earlier investigations have revealed the general principles of helix-to-helix packing in globular proteins (32). Exploring these principles, Cohen *et al.* (33) developed a method for the generation of three-dimensional protein structures from the secondary structure assignment.[2] These methods have been applied to myoglobin, tobacco mosaic virus coat protein, growth

* This work was supported by National Institutes of Health Grants RO1-HL42949 (to H. F. B.) and RO1-GM 39900 (to F. E. C) and by a grant from the R. W. Johnson Pharmaceutical Research Institute. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] The abbreviations used are: Epo, erythropoietin; IL, interleukin; bp, base pair(s); RIA, radioimmunoassay; IPTG, isopropyl-1-thio-β-D-galactopyranoside; WT, wild type.

[2] These algorithms are available from Dr. Cohen upon request.

hormone, α- and β-interferon, IL-2, and IL-4 (33–36).

The algorithm for tertiary structure generation is divided into four computations. The program *aapatch* identifies clusters of hydrophobic residues within the putative helices that could mediate helix-helix interactions (32). *Aafold* generates all possible helix pairings according to the location and geometric preferences of the interaction sites. *Aabuild* generates the three-dimensional models of all possible structures from the list of helix pairings (from *aafold*) and subject to steric restrictions and geometric constraints on chain folding. In the final step, *aavector* applies the user-defined distance constraints (*e.g.* disulfide bridges) to the structures generated. At this stage, coordinates have been specified only for residues in the core α-helices. For residues in sequentially distinct loops, lower bounds on the inter-residue distances can be inferred from the relevant helix terminus.

Preparation of Epo Muteins

Construction of the Mutagenic/Mammalian Expression Plasmid—A M13 plasmid, containing a 1.4-kilobase *Eco*RI-*Eco*RI human Epo cDNA insert (λHEPO FL12) was a gift from Genetics Institute (Cambridge, MA) (8). A 943-bp *Eco*RI-*Bgl*II fragment, corresponding to the complete coding sequence of the wild type human erythropoietin, including untranslated regions 216 bp upstream and 183 bp downstream, was inserted into the mammalian expression plasmid pSG5 (Stratagene) (37) and named pSG5-EPO/WT.

Site-directed Mutagenesis— was carried out according to the protocol described by Kunkel *et al.* (38). Single-stranded DNA was rescued from the pSG5-EPO/WT phagemid grown overnight in *Escherichia coli* CJ236, in 2XYT media containing M13KO7 helper phage (In Vitrogen) and 70 μg/ml kanamacyn (Sigma). The resulting uracil-containing single-stranded DNA was used as a template for mutagenesis. Oligonucleotides (24–46-mer) were synthesized with their 5′ and 3′ ends complementary to the target wild type Epo sequence. A large variety of mutations (base substitutions, deletions and insertions) were created at the centers of the mutagenic primer sequences. Annealing of the phosphorylated primers (10:1 oligonucleotide/DNA template molecular ratio) was performed in 10 μl of a 20 mM Tris-HCl, pH 7.4, 2 mM $MgCl_2$, 50 mM NaCl solution. The reactions were incubated at 80 °C for 5 min and then allowed to cool slowly to room temperature over a 1-h period. The DNA polymerization was initiated by the addition as a mix of 1 μl of 10 × synthesis buffer (100 mM Tris-HCl, pH 7.4, 50 mM $MgCl_2$, 10 mM ATP, 5 mM each dNTPs, 20 mM dithiothreitol), 0.5 μl (8 units) of T4 DNA ligase and 1 μl (1 unit) of T4 DNA polymerase (Boehringer Mannheim). After 2 h at 37 °C, 80 μl of 1 × Tris-EDTA was added. 5 μl of the diluted reaction mix was used to transform competent *E. coli* NM522 (ung^+, dut^+).

Since a 40–80% mutation yield is normally obtained, four to five double-stranded plasmid clones from each reaction were sequenced with 7-deaza-dGTP and Sequenase (U. S. Biochemicals Inc.) (39). As a rule, the entire coding sequences of the Epo mutants were examined for the presence of unwanted mutation by sequencing or restriction enzyme mapping.

Production of Wild Type and Epo Muteins in Mammalian Cells—Cos7 cells grown to ~70% confluence were transfected with 10 μg of recombinant plasmid DNA/10-cm dish using the calcium phosphate precipitation protocol (40). As a control of transfection efficiency, in several experiments 2 μg of pCH110 plasmid (Pharmacia LKB Biotechnology Inc.) was cotransfected and β-galactosidase activity measured in the cytoplasmic extracts.

RNA Blot-hybridization Analysis—Total RNAs were prepared from cultured Cos7 cells (41) and 2-μg samples electrophoresed on a 1.1% agarose gel containing 2.2 M formaldehyde. Transfer to GeneScreen Plus filters (Du Pont-New England Nuclear) and hybridization with a ^{32}P-labeled WT Epo probe were carried out as previously described (42).

Quantitation of Transiently Expressed Recombinant Epos—The amount of secreted protein in the supernatants of transfected Cos7 was determined by a radioimmunoassay (RIA). The RIA was performed using a high titer rabbit polyclonal antiserum raised against the human wild type Epo and produced in our laboratory. ^{125}I-Labeled recombinant Epo was obtained from Amersham Corp. Details of the protocol have been published elsewhere (43).

Immunoprecipitation of ^{35}S-Labeled Epo Proteins—Three days after transfection, the Cos7 monolayers were washed extensively with 1 × phosphate-buffered saline and the cells incubated for 20 min at 37 °C in 2 ml of Met^-/Cys^- minimum essential media Eagle's modified medium. In each culture dish, 100 μl of TRAN^{35}S-LABEL ([^{35}S] cysteine and [^{35}S]methionine, ~10 mCi/ml, ICN Biochemical) was then added. After 2 h, the conditioned media were harvested, and cellular extracts were prepared by lysis in radioimmune precipitation buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 0.02% (w/v) sodium azide, 0.1% (w/v) SDS, 0.5% (w/v) sodium deoxycholate, 1% (w/v) Triton X-100, 1 mM phenylmethylsulfonyl fluoride, and 1 μg/ml aprotinin). Samples were precleared with rabbit preimmune serum/protein A-Sepharose CL-4B (Pharmacia) for 2 h. Immunoprecipitations were performed overnight with our polyclonal antibody specific for human recombinant wild type Epo and immunoadsorbed with protein A-Sepharose CL-4B. Immunoprecipitates were run on 15% SDS-polyacrylamide gels (44) and analyzed by autoradiography after treatment with Enhance (Du Pont-New England Nuclear).

Bioassays—The dose-dependent proliferation activities of WT and Epo muteins were assayed *in vitro* using three different target cells: murine spleen cells, following a modification of the method of Krystal (45, 46); murine Epo-responsive MEL cell line, developed by Hankins (47); and human Epo-dependent UT-7/Epo cell line, derived from the bone marrow of a patient with acute megakaryoblastic leukemia (48). After 22–72 h of incubation with increasing amounts of recombinant proteins, cellular growth was determined by [^3H]thymidine (Du Pont-New England Nuclear) uptake or by the colorimetric MTT assay (Sigma) (49).

Bacterial Expression—The wild type Epo target, corresponding to the nucleotide sequence coding for the mature protein, was polymerase chain reaction-amplified using appropriate primers. In the sense primer an *NdeI* site (CATATG) was placed immediately 5′ to the Ala^1 codon of the mature protein. In the antisense primer a *Bgl*II site was placed 3′ to the TGA stop codon. After enzymatic digestion, the 516-bp polymerase chain reaction fragment was inserted in an *NdeI*/*Bam*HI-cut pET16b plasmid (Novagen), which has a T7 promoter followed immediately by the lac operator. IPTG induction of transformed *E. coli* BL21(DE3)(T7 RNA polymerase$^+$, lon^-, $ompT^-$) resulted in high levels of expression of a fusion protein with a 10-histidine stretch at the amino terminus. The oligo-His tag allowed the binding of the produced (His_{10})-Epo on a nickel affinity resin and its elution by increasing imidazole concentrations in presence of phenylmethylsulfonyl fluoride (Sigma). Most of the produced protein formed insoluble aggregates and was solubilized and affinity-purified under denaturing conditions in 6 M guanidine HCl. Oxidative refolding was performed by overnight dialysis at 4 °C against 50 mM Tris-HCl, pH 8.0, 40 μM $CuSO_4$, and 2% (weight/volume) Sarkosyl. Soluble protein was further dialyzed against 20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 2 mM $CaCl_2$, and subjected to factor Xa (New England Biolabs) cleavage to remove the NH_2-polyHis sequence. Monitoring of the fusion protein following induction and during the various steps of purification was done by electrophoresis on a 15% polyacrylamide-SDS gel, stained with Coomassie Brilliant Blue. Alternatively, the His-Epo fusion protein was detected on Western blot (50), using a 1/2000 dilution of our WT native Epo polyclonal antibody and a second biotinylated rabbit-specific antibody which is detected with a streptavidin-alkaline phosphatase conjugate (Amersham).

In Vitro Transcription/Translation—Sense and antisense primers, creating new *Bgl*II sites, respectively, 5′ and 3′ of the initiator and stop codons, were used in a polymerase chain reaction on pSG5-EPO/WT template. After *Bgl*II cleavage, the 594-bp polymerase chain reaction fragment was subcloned into pSP64T (51). This SP6-containing vector provides 5′- and 3′-flanking regions from *Xenopus* β-globin mRNA, which allow efficient *in vitro* transcription/translation. Previous experiments showed poor yields of *in vitro* translated protein, when using the GC-rich natural 5′ Epo untranslated region. One step *in vitro* transcription/translation was carried out by incubation of 1 μg of circular p64T-Epo in a 50-μl reaction volume of SP6-TnT-coupled rabbit reticulocyte lysate system (Promega), in the presence of [^{35}S]cysteine (1200 Ci/mM, Du Pont-New England Nuclear). In some cases, canine pancreatic microsomal membranes were added to the reaction mix. A purified GST-human Epo receptor extracellular domain fusion protein (EREx) was a gift from W. Harris and J. Winkelman, and the binding of the ^{35}S-labeled translation products onto EREx-glutathione agarose beads was performed as described (52).

RESULTS

Construction of a Model of the Three-dimensional Structure of Erythropoietin

From an analysis of the putative Epo helix sequences, *aapatch* identified eight possible helix-helix interaction sites.

TABLE I

Predicted α-helical regions of the mature erythropoietin protein

Data were obtained using the various algorithms for secondary and tertiary structure generations described under "Materials and Methods."

Helix	NH_2 terminus	COOH terminus	Potential helix-helix interaction sites
A	9	22	19
B	59	76	63, 67, 70, 71
C	90	107	95, 102
D	132	152	141

In principle, these sites could be used to generate 1.6×10^4 structures. Of these, only 706 maintained the connectivity of the chain and were sterically sensible. These structures resembled four helix bundles, an increasingly common motif in protein structure (53). The structures that were not compatible with the native disulfide bridge between Cys_7 and Cys_{161} were eliminated. This reduced the total number of structures from 706 to 184 (total computer time approximately 1 h on a Silicon Graphics IRIS 4D/35G). The remaining structures were then rank ordered by solvent-accessible surface contact area, a measure of the validity of model structures. The most compact structures were right handed, all anti-parallel four-helix bundles with no overhand connections, but this may be an artifact of a failure to add the polypeptide chain that forms the loops to the helical core constructed by *aabuild*. The other less compact structures were left-handed four-helix bundles with two overhand loops, a topology previously seen in the structures of IL-4 and growth hormone. We suspect that this is the likely structure for Epo. The consensus for assignments of putative α-helices in human Epo are summarized in Table I. First, analysis of the topological distribution of known four-helix bundle structures indicates that nearly all examples have an antiparallel orientation (53). Second, the left-handed four-helix bundles with two overhand connections arrange the four amphipathic helices to form a compact hydrophobic core. Finally, the AB and CD loop regions of Epo are predicted to have β-sheet segments analogous to IL-4 and growth hormone that preserves the compact globular nature of the Epo model structure. Fig. 1 shows schematic representations of predicted topological interactions between the four anti-parallel α-helical bundles.

Several authors have suggested that the helical cytokines form a structural superfamily (34, 54–57). On the basis of both the mature protein and the individual α-helices, Epo seems to be more closely related to growth hormone, prolactin, IL-6, and GM-CSF rather than the other members of the helical cytokine superfamily. Nevertheless, recent improvements in algorithms for the identification of distant evolutionary relationships between proteins from structural fingerprints suggested that it might be possible to align the IL-4 structure to the Epo sequences. The Eisenberg *et al.* (58) structural environment and 3D-1D profile methods are a powerful tool for recognizing that a sequence is compatible with a known structure, *e.g.* a four-helix bundle. The NMR structure of IL-4 from Smith *et al.* (59) was used to construct a 3D-1D profile. A mixture of sequences including four helix bundles, globins, and non-helical structures were aligned against the IL-4 profile. Not surprisingly, the IL-4 structures from human and mouse gave the highest scores (Z[3] = 22.8 and 8.1). However, the other known four-helix bundle cytokines known to share a similar fold with IL-4, *e.g.* human growth hormone (60) (Z = 2.3) and GM-CSF (61) (Z = 2.3) faired no better than some globin sequences (Kuroda's and slug sea hare globin, Z = 5.0 and 4.8) that adopt a distinct tertiary structure. The results for the human and sheep Epo sequences were also ambiguous (Z = 1.6 and 0.8). These results suggest that while profile methods are a powerful tool for recognizing structural similarity, their failure to identify homology does not exclude the possibility that two proteins share a common fold. For distantly related or unrelated structures, current profile methods cannot replace *de novo* methods for tertiary structure prediction.

[3] Z scores are used to describe the normalized weight associated with a profile score. A distribution is built from a collection of sequences with a mean Z score of 0.0 and a standard deviation of 1.0. Z scores greater than 6.0 are associated with significant alignments. Z scores between 3.0 and 6.0 may or may not be structurally relevant.

Design and Expression of Epo Muteins That Test the Proposed Structure

To test the proposed four α-helical bundle structure of erythropoietin and at the same time to attempt to locate functional domains, we created by site-directed mutagenesis a series of deletion, insertion, and replacement mutants. These muteins were designed to analyze the principal predicted structural features of the molecule: α-helices, interconnecting loops, as well as the NH_2 and COOH termini. Structural and functional implications of the disulfide bridges and the glycosylation sites were also investigated.

α-Helices—Short amino acid deletions were prepared in, or close to, the predicted A, B, C, and D α-helices. Human wild type and muteins were transiently expressed in Cos7 cells. Northern blot analyses demonstrated that all the mutant plasmids produced about the same amount of mRNA as that of the wild type (data not shown). Yet, no detectable amount of Epo protein could be found in the Cos7 supernatants, either by radioimmunoassay or by bioassay using various Epo-dependent cell lines. Table II summarizes these findings.

An example of SDS-polyacrylamide gel electrophoresis of immunoprecipitants from *in vivo* ^{35}S labeling is presented in Fig. 2. As expected, when Cos7 cells were transfected with pSG5-EPO/WT, a 35–37 kDa band was detected in the supernatant. In contrast, the deletion mutants (Table II) could be detected in cellular extracts but were not exported from the cells. Fig. 2 shows the cytoplasmic retention of the mutein Δ140–144, lacking 4 residues in the middle of the predicted D-helix. The apparent molecular mass (~ 28 kDa) is less than expected for a 5-amino acid deletion. Therefore, not only the secretion, but also the glycosylation, seem to be impaired. None of these muteins had deletion of glycosylation sites. It is likely that full glycosylation of Epo requires conservation of its molecular architecture. Similar results (reported in Table II) were obtained for all the muteins having partial deletion of an α-helical peptide segment.

Because contaminants in crude Cos7 cellular extracts severely interfere with the radioimmunoassay, no direct Epo quantitation was possible. However, aliquots of hypotonic extracts of Cos7 transfected with wild type Epo were able to sustain HDC57 proliferation. No similar biological activity was found for muteins with limited deletion of α-helices.

Interconnecting Loops—The peptide segment joining A- and B-helices presents several interesting features (Fig. 3*A*). AB loop consists of 36 amino acids. Two *N*-glycosylation sites and a small disulfide bridge are located in the first half and their biological implications will be discussed later. The COOH end of the AB loop contains a stretch of amino acids that is strongly conserved among mammals (26). Alignments of human, monkeys, cat, mouse, rat, pig, and sheep Epos showed a consensus sequence: DTKVNFYAWKR(M/I)(E/D)VG (residues 43–57). Three deletions were constructed:

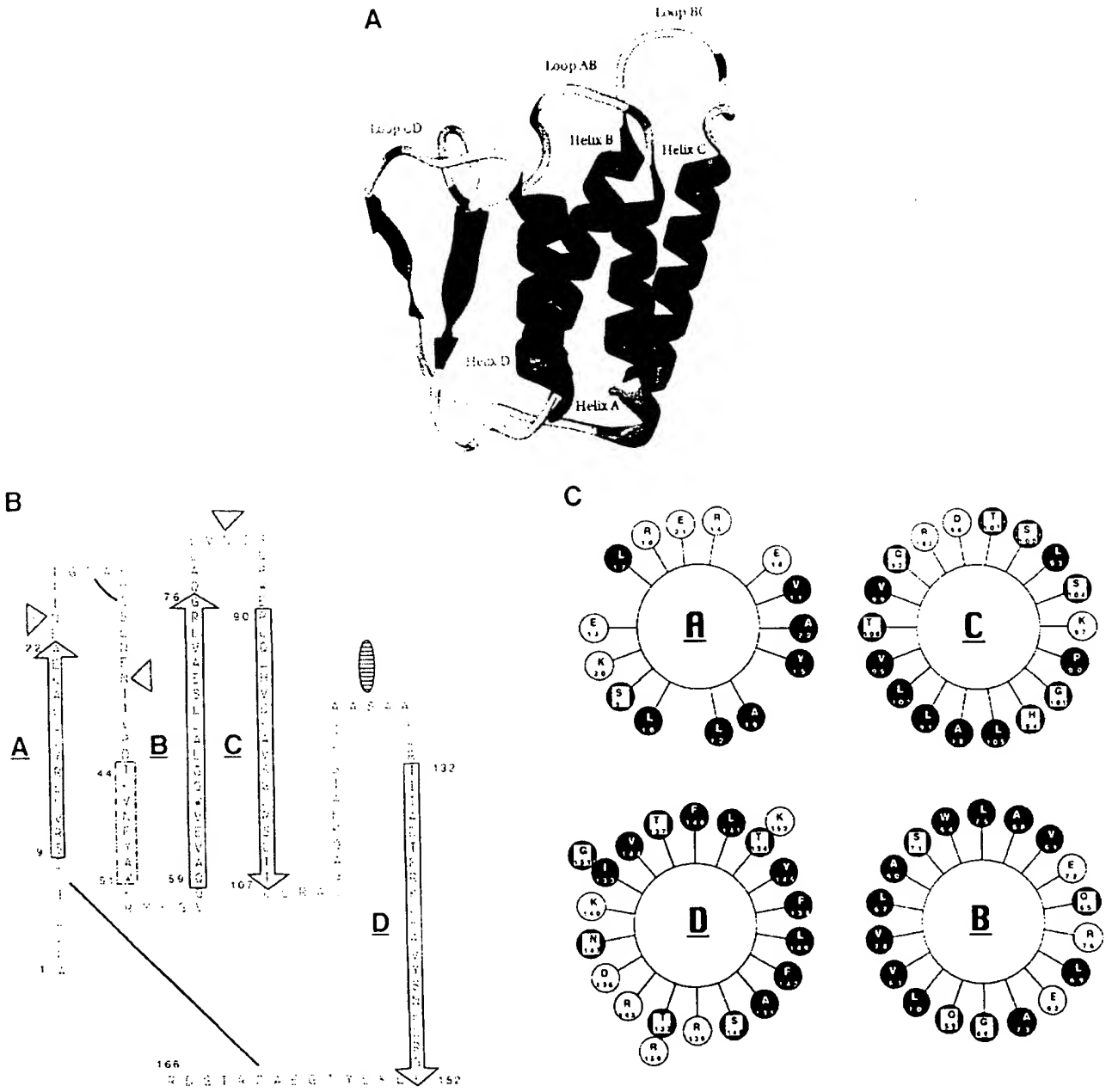


FIG. 1. **Model of the three-dimensional structure of erythropoietin.** *A*, ribbon diagram of the predicted Epo tertiary structure. The four α-helices are labeled *A–D* (*magenta*); Loops between helices are named for the helices they interconnect. Two regions of extended structure which could form hydrogen bonds between *Loop AB* and *Loop CD* are also presented (*cyan*). *N*- and *O*-glycosylation sites are indicated in *green* and *blue*, respectively. Disulfide bonds bridge residues 29–33 in *Loop AB*, and 7–161 on the NH_2-terminal side of *Helix A* and the COOH-terminal side of *Helix D* are not shown. *N.B.*: The loop tracing shown does not represent predicted coordinates. *B*, schematic representation of Epo's primary structure depicting predicted up-up-down-down orientation of the four antiparallel α-helices (*boxes* with *arrowhead*). This folding pattern is strongly suggested by the large size of the two interconnecting loops *AB* and *CD*. The limits of each helix were drawn accordingly to Table I. A predicted short region of β-sheet is delineated by the *dashed rectangle*. The *N*-glycosylation sites are represented by the *dotted diamonds*, and the *O*-glycosylation site by the *dashed oval*. The locations of the two disulfide bridges are shown. *C*, cross-section of the Epo molecule at the level of the four α-helices. The helical wheel projections are viewed from the NH_2 end of each helix. The hydrophobic residues, localized inside the globular structure, are indicated by *filled circles*. The charged and neutral residues (*open* and *gray circles*, respectively) are exposed at the surface of the molecule.

Δ43-47, Δ48-52, and Δ53-57, and transiently expressed in Cos7. The amount of muteins detected by RIA in the supernatants of transfected cells was 10–40% lower than observed with wild type Epo (Fig. 3*B*). Nevertheless, the three secreted muteins were biologically active. However, because Δ48–52 exhibited a marked decrease of the specific bioactivity, this site was studied in more detail by means of serine replacements. Krystal *ex vivo* bioassay as well as HCD57 and UT7-Epo *in vitro* bioassays showed that these Ser mutants had biological activities similar to that of wild type (Fig. 3*C*). Therefore, the observed decreases in both RIA and bioassay for the three deletion mutants are likely to be the result of changes of structural conformation. The long length of loop AB may be critical for the up-up-down-down topography. A

TABLE II

Short deletions in, or close to, α-helices

Predicted NH_2 and COOH termini of each α-helix are indicated in the vertical boxes.

Mutants			
Δ12–16	9	A helix	
	↓		
	22		
Δ65–69	59	B helix	RNA levels comparable to WT.
	↓		
	76		
Δ96–100	90	C helix	No detectable Epo in the Cos7 supernatant, both by RIA and bioassay.
Δ105–109	↓		
	107		
Δ122–126			
Δ131–135	132	D helix	
Δ140–144	↓		
Δ142–150			
Δ152–155	152		
Δ156–160			

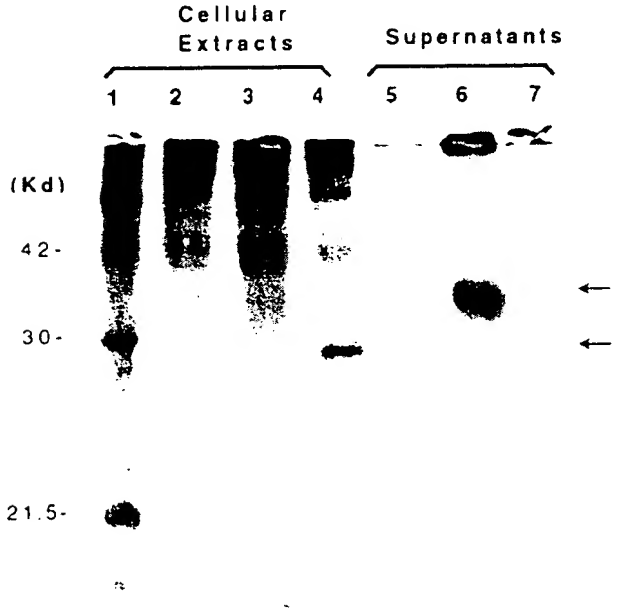


FIG. 2. **Immunoprecipitations of wild type Epo and the Δ140–144 mutein.** Cos7 cells were transfected with pSG5, pSG5-EPO/wt, or pSG5-EPO/Δ140–144. After 3 days, the cells were metabolically labeled with [^{35}S]methionine and [^{35}S]cysteine. Immunoprecipitations of cellular extracts and supernatants were performed with our polyclonal antibody, raised in rabbit against the native human Epo. The immunoprecipitates were analyzed by SDS-polyacrylamide gel electrophoresis. *Lanes 2–4* correspond to the cellular extracts; *lanes 5–7* in the culture supernatants from transformed Cos7 with: plasmid without insert (*lanes 2* and *5*), wild type Epo (*lanes 3* and *6*), and Δ140–144 (*lanes 4* and *7*). *Lane 1* represents the protein molecular weight standard. The *two arrows* show the normal secretion of the wild type Epo (35–37 kDa) and the cytoplasmic retention of the mutein Δ140–144 (~28 kDa).

shorter AB segment may impose a strain on the interhelical connection. Chou and Fasman (30) algorithms predicted a short β-sheet structure from residues 44 to 51 (<Pa>;<Pb>, 1.005 < 1.196). The presence of a short region of β-sheet in the connection between helices 1 and 2 (A and B) have been documented in the analyses of the three-dimensional structures of IL-4 (22, 23), GM-CSF (24), and monomeric M-CSF (25). In contrast, in human GH a short segment of α-helix is found at the same location (20). The structure/function implications of these short features are not yet understood.

Helix B is linked to helix C by a much shorter segment (residues 77–89) and contains in its center the third *N*-glycosylation site (Asn^{83}). When the Δ78–82 mutein was expressed, a secreted protein was detected in the conditioned medium and conferred proliferative bioactivity on Epo-dependent cell lines (see Fig. 8).

A similar long crossover connection (23 amino acids) is found between helix C and helix D. In contrast to what we previously observed for loop AB, a large deletion of 9 residues at position 111–119 or a 7-amino-acid insertion of a myc epitope after residue 116 did not affect the secretion of these muteins (Fig. 4). Furthermore, these two proteins had normal specific activity, as seen by the ratio of bioassay to RIA. Our rabbit polyclonal antibody raised against the native form of the human wild type fully recognized the two mutants, demonstrating that the overall spatial conformation of Epo was well preserved. According to the algorithm of Emini *et al.* (62), the residues 111–119 are predicted to be at the surface of the molecule. Since the Δ111–119 mutein is readily secreted and has full biological activity, it seems unlikely that the putative β-sheet segment in the CD loop is an important determinant of molecular stability. Primary amino acid alignments of mammalian Epo showed a large variation in the sequence of residues 116–130, including amino acid deletion, insertion, and substitution (26). Surprisingly, when the deletion Δ122–126 mutein, which removed the *O*-glycosylation site (Ser^{126}), was transiently expressed in monkey cells, protein secretion was inhibited. Both rodents, rat and mouse, lack the *O*-glycosylation site because of a Ser^{126} to Pro replacement. Furthermore, when a Ser^{126} replacement mutein was expressed in normal Chinese hamster ovary cells, (63) or when wild type Epo was expressed in cells having a defect in *O*-linked glycosylation (64), neither secretion nor biological activity were impaired. Therefore, failure of secretion of the Δ122–126 mutein may be the result of some other structural alteration. In particular, the proline residue at position 122 is invariant among mammals.

NH_2 and COOH Termini—Deletion of residues 2 to 5 only slightly affected the processing of a biologically active protein (see Fig. 8). This deletion may impair cleavage of the propeptide, therefore explaining the lower yield of secreted Epo mutein in comparison to that of wild type. The fact that the mature monkey protein has an elongated (Val-Pro-Gly) NH_2 terminus strongly suggests that the NH_2-terminal part is not involved in the bioactivity of the molecule. Further evidence comes from the results, reported below, on the *N*-poly-His-Epo fusion protein expressed in *E. coli*, and also from the identical binding of *in vitro* translated ^{35}S-labeled wild type Epo onto EREX-glutathione agarose beads (Fig. 5), with or without addition of canine pancreatic microsomal membranes which permit cleavage of the propeptide.[4]

The COOH-terminal sequence following helix D can clearly be divided into two distinct domains, separated by Cys^{161}. The residues 151–161 were of special interest because they are highly conserved among mammals (26). There are only two substitutions: Lys^{164} is replaced by a Thr in artiodactyls and cat, and Ala^{160} is replaced by a Val in mouse Epo. Both the Δ152–155 and the Δ156–160 muteins remained in the cytosol of the transfected Cos7 (Table II). One possible explanation is that the residues 152–160 may, in fact, participate in the D helix. We predict that Gly^{151} is the break point of the struc-

[4] All the mutants described in this paper were subcloned into pSPG4T plasmid. Studies of the binding of their translation products to EREx are in process.

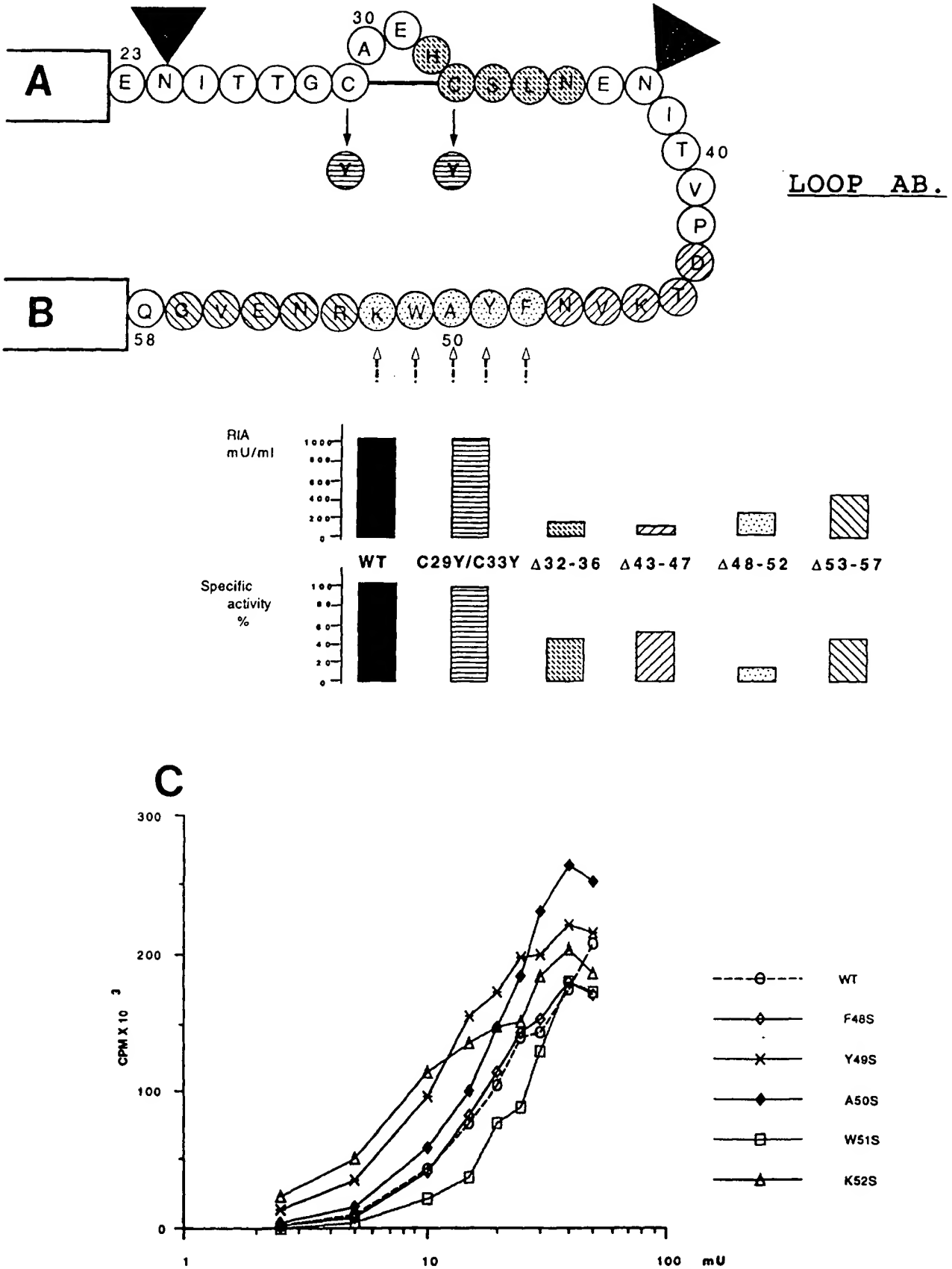


FIG. 3. **Interconnecting loop AB.** *A*, schematic representation of the loop AB showing the localization of muteins with various deletions and amino acid replacaements. The *dashed arrows* point to the positions of the serine substitutions (in Δ48–52). The two *N*-glycosylation sites are represented by the *gray diamonds*. The small Cys^{29}=Cys^{33} disulfide bridge is indicated. *B*, amount and biological activities of secreted

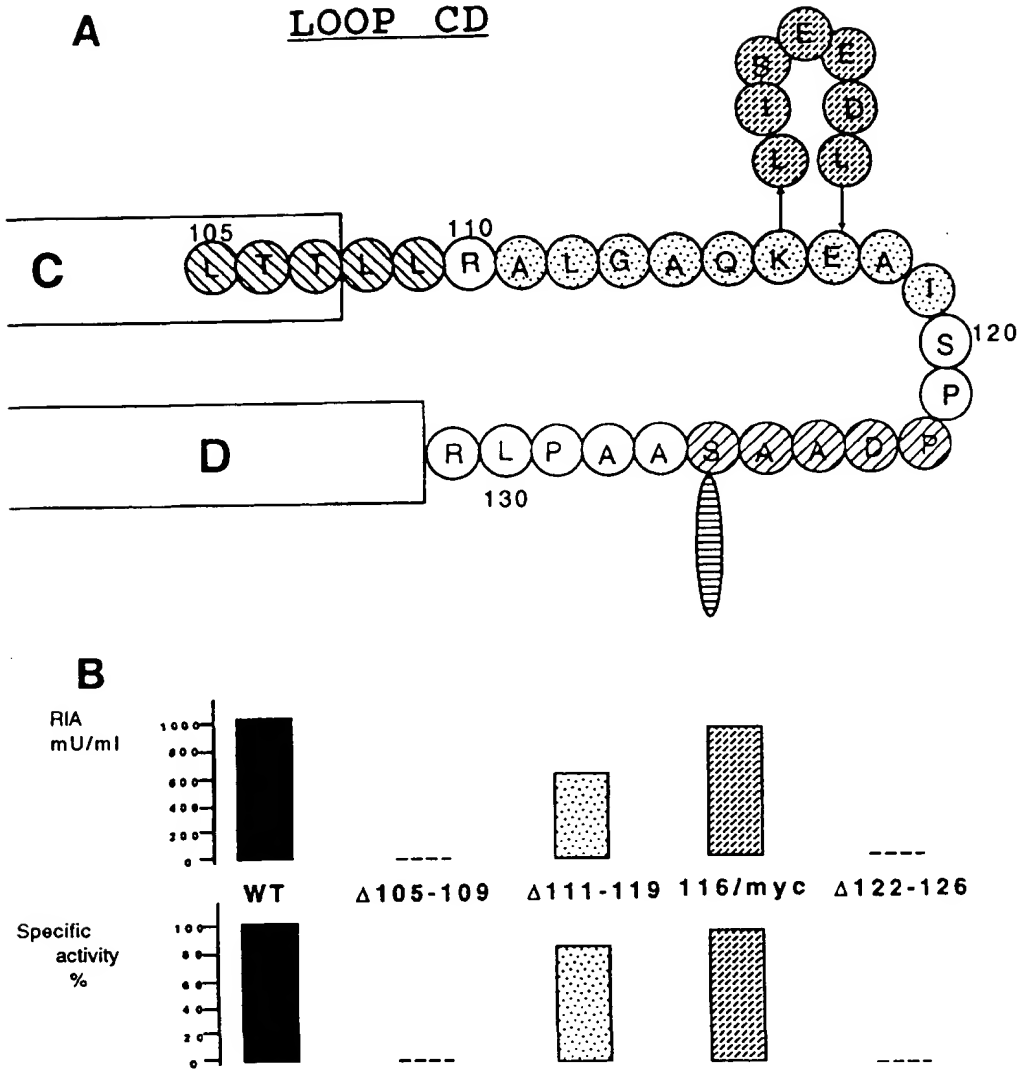


FIG. 4. Interconnecting loop CD. *A*, schematic representation of the loop CD showing the location of three deletion muteins: Δ105–109, Δ111–119, Δ122–126, and the insertion of 7 residues after Lys*166* (myc epitope). The *O*-glycosylation site is indicated by the *dashed oval*. *B*, secretion and biological activities of the muteins located in loop CD. The *two bar graphs* were created as described in Fig. 3*B*. The two mutants Δ111–119 and 116/myc were normally secreted and had full biological activities. *mU*, milliunit.

ture. However, it is possible that this residue causes only a bend in the α-helical structure and helix D may extend to Gly[158]. The COOH-terminal part of the protein (residues 162–166) is clearly not involved in any structural or functional feature. Thus, the deletion of the 4 last amino acids or the replacement of residues 162–166 by either a KDEL sequence or a poly-histidine sequence[6] did not modify the specific activity of the erythropoietin (Fig. 6). Radioimmunoassay revealed that the secretion of the KDEL-tail mutein in the media of transfected cells was 45% less than normally obtained with the wild type Epo. However, when compared to the wild type, this mutein had more biological activity in the hypotonic Cos7 cell extracts. The KDEL COOH-terminal sequence has been shown to be essential for the retention of several proteins in the lumen of the endoplasmic reticulum (65). Nevertheless, because of overproduction in transiently expressed cells, a large percentage of recombinant protein escaped into the media.

Disulfide Bridges—Wang *et al.* (66) demonstrated that the biological activity of Epo was lost irreversibly if the sulfhydryl

[6] The poly-His tail wild type mutant was purified by means of nickel affinity chromatography which enables quantitation of cytosolic-retained mutants. The $(His)_6$ COOH-terminal sequence has been appended to all the muteins described in this paper. Experiments are in progress to exploit this strategy.

muteins. The *upper bar graphs* show the relative secretion of wild type and loop AB muteins as determined by radioimmunoassay. The *lowest bar graphs* display the calculated specific activity (ratio bioassay/RIA) for each mutein, in comparison with the value obtained for the wild type Epo (ratio = 100%). *C*, HCD57 cell proliferation as a function of increasing concentration of wild type and serine-substituted Epo muteins. HCD57 cells (10^4/ml) were cultured for 3 days in a 96-well microtiter plate with media containing increasing concentrations of secreted proteins. The *line graphs* show the cellular growth as measured by [^{3}H]thymidine uptake. The number of viable cells was also measured with the MTT colorimetric assay and gave similar curves. *In vitro* proliferation experiments using the human UT-7 cell line (48) and the Krystal assay (45) produced identical results. *cpm*, counts/minute; *mU*, milliunit.

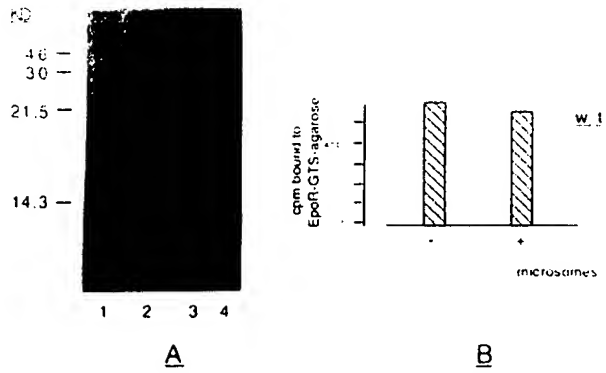


FIG. 5. *In vitro* translation of the Epo wild type. *A*, analysis of the ^{35}S-labeled translation products by SDS-polyacrylamide gel electrophoresis. One-step transcription/translation reactions were performed in the SP6-TnT rabbit reticulocyte lysate system. 1/30 of each reaction was resolved on a 15% polyacrylamide gel. *Lane 1*, low *M*, standard from Amersham Corp.; *lane 2*, *in vitro* reaction without added plasmid; *lanes 3* and *4*, translation products obtained after incubation of 1 μg of circular p64T-Epo, respectively, in the presence or absence of canine pancreatic microsomal membranes. *B*, binding of the *in vitro* translated Epo wild type onto Epo receptor-GTS-agarose beads. 6×10^3 count/min of purified ^{35}S-labeled erythropoietin products, processed with microsomes (+) or not (−) were incubated in the presence of EREx, following the protocol described by Harris *et al.* (52). Identical binding demonstrated that the conservation of the propeptide did not impair the hormone-receptor interaction.

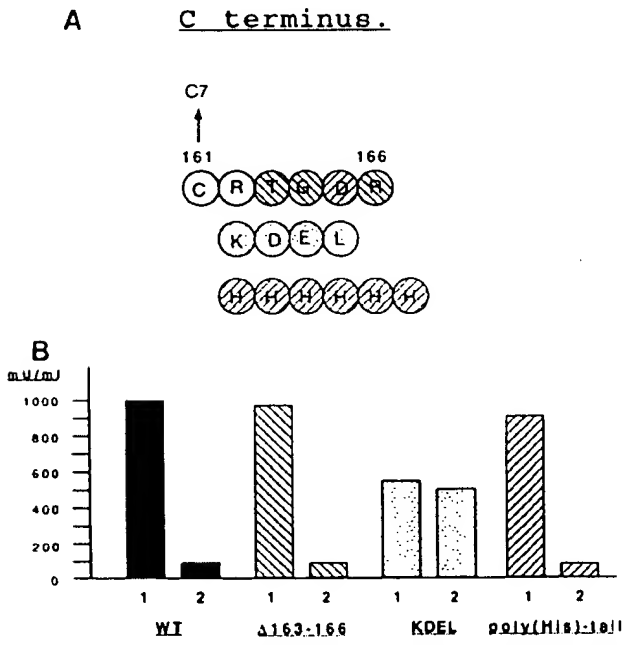


FIG. 6. COOH end of Epo. *A*, schematic representation of the analyzed muteins, corresponding to the deletion of the four last amino acids Δ163–166 and the replacements of the residues 162–166 by a KDEL or poly (His) sequences. *B*, relative secretion of these muteins. The bioactivities in the supernatants (*1*) and the cell extracts (2) of transformed Cos7 cells were measured by *in vitro* proliferation assay using HCD57. More KDEL mutant remained in the cytosol of the Cos7, when compared with the wild type Epo and Δ163–166 or poly (His) muteins. However, all the analyzed muteins had the same specific activity as that of the wild type. *mU*, milliunit.

groups were alkylated. The mature human Epo has two internal disulfide bonds: Cys^7-Cys^{161}, linking the NH_2 and COOH termini of the protein and a small bridge between Cys^{29} and Cys^{33}. Cys^{33} was previously changed to Pro by site-directed mutagenesis, and the resulting protein was reported to have greatly reduced *in vitro* biological activity (67). However, rat and mouse Epos have the same substitution and yet exhibit full cross-species bioactivity. To resolve the role of the small disulfide bridge in human Epo function, we created a C29Y/C33Y double mutation. The resulting mutein was normally processed and showed the same *in vitro* bioactivity as the wild type (Fig. 3*B*).[6] Furthermore, the deletion of 5 amino acid residues (Δ32–36) did not impair the secretion of a biologically active mutein. These data suggest that only the native and fully conserved disulfide bridge Cys^7-Cys^{161} is crucial for the preservation of the molecular structure of erythropoietin.

Functional Role of the Glycosylation—Natural or recombinant human Epo is a heavily glycosylated protein; 40% of its molecular weight is sugars (11). The protein has three *N*-linked oligosaccharide chains, located at amino acid positions 24 and 38 (in predicted loop AB) and position 83 (in loop BC). It has one *O*-linked carbohydrate chain at position 126 (in loop CD), which is missing in rodents. The role of these sugar chains in the biological activity of the human hormone has been extensively studied. Site-directed mutagenesis at the *N*-glycosylation sites demonstrated that even though the sugars were important for proper biosynthesis and secretion, their removal did not affect *in vitro* activity. This finding was corroborated by several investigators (68, 69). However, Takeuchi *et al.* (70) found that *N*-glycanase digestion results in almost complete loss of biological activity. In contrast, there is general consensus that glycosylation plays a key role in the biological activity of the hormone *in vivo*. Various reports have demonstrated that the *N*-linked sugar chains enhance the stability and survival of Epo in the blood stream (71, 72) and protect the hormone against clearance by the liver (73), thereby enabling the transit of the hormone from its site of production in the kidney to its target cells in the bone marrow (74).

We expressed the wild type Epo in *E. coli*. Accordingly, the produced protein completely lacks sugar. The pET expression system was used and is detailed under "Materials and Methods." IPTG induction of transformed BL21(DE3) bacterial strain rapidly results in a high level of expression of the poly-His Epo fusion protein (Fig. 7, *A* and *B*). After 3 h of induction, we obtained a typical yield of ~1 mg of total protein/ml of culture. However, the vast majority of the expressed protein was present in the inclusion bodies, and therefore its solubilization and purification on the nickel beads were performed in 6 M guanidine HCl. Oxidative refolding and factor Xa cleavage resulted in soluble forms (Fig. 7*C*), and the *in vitro* biological activity was tested on HDC57 cells. The cleaved *E. coli* recombinant Epo showed a notable decrease of the specific activity (10% less than the fully glycosylated mammalian expressed protein), but was still able to maintain HCD57 proliferation. The observed reduction of *in vitro* activity is likely to be due to improper refolding of the insoluble protein and impaired physical stability of the *E. coli* Epo as previously reported (55). However, the fact that the *E. coli* Epo was still able to trigger HCD57 growth indicated an overall preservation of the molecular structure. The uncleaved fusion protein, with 10 His residues at the NH_2 terminus, exhibited a 67% loss in biological activity when compared to the cleaved protein. Thus, the addition of a 20-residue sequence to the NH_2 terminus partially inhibited the biological activity.

[6] Two single replacement muteins (C29Y and C33Y) were also stable and had full biological activity (results not shown).

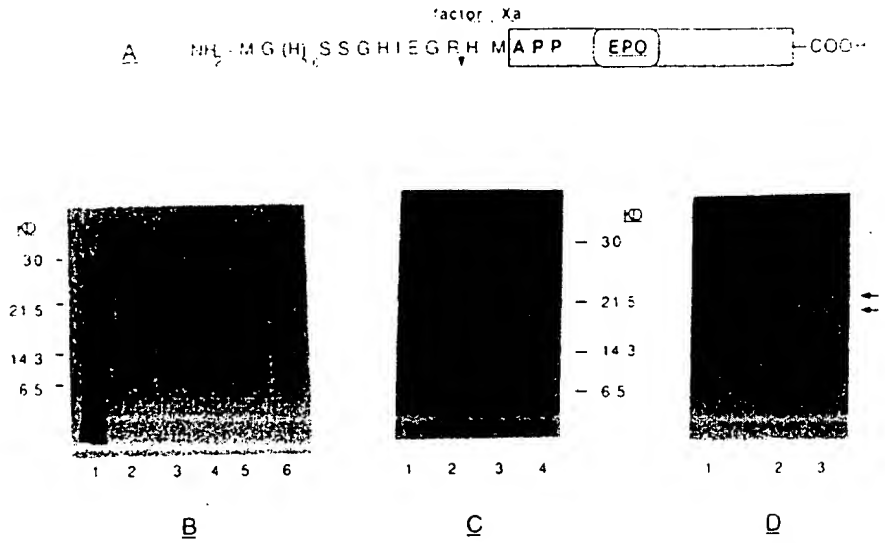


FIG. 7. **Bacterial expression of wild type Epo.** *A*, diagram of the fusion protein. An NH_2-terminal 22-amino-acid long peptide, containing a 10 histidine stretch, was fused to the mature erythropoietin sequence. Factor Xa cleavage allowed the recovery of the mature Epo with only 2 extra residues at its amino terminus. *B*, IPTG induction of the fusion protein. Transformed *E. coli* BL21(DE3) cultures (OD_{600} = 0.6) were grown in the presence of 1 mM IPTG. Aliquots were collected at 0 (*lane 2*), 1 (*lane 3*), 2 (*lane 4*), and 3 h (*lane 5*) and analyzed on a 15% SDS-polyacrylamide gel, stained with Coomassie Brilliant Blue. A high level of production of the fusion protein was rapidly obtained. *Lane 6* corresponds to an aliquot from transformed bacteria grown for 3 h in a medium without IPTG. *Lane 1* is a low molecular weight standard. *C*, purification of the fusion protein. After 3 h of IPTG induction, the produced $(His)_{10}$-Epo was solubilized in 6 M guanidine HCL and purified on a nickel affinity resin by increasing imidazole concentrations following the pET-His system protocol (Novagen). Samples of the column eluants were analyzed by SDS-polyacrylamide gel electrophoresis and stained with Coomassie Brilliant Blue. *Lane 1*, elution by 20 mM imidazole; *lane 2*, elution by 100 mM imidazole, releasing the fusion protein; *lane 3*, chelation of the nickel by a 100 mM EDTA wash; *lane 4*, molecular weight standard. *D*, detection of the *E. coli* recombinant Epo on a Western blot. Solubilized proteins were separated on a 15% SDS-polyacrylamide gel, transferred to a nitrocellulose membrane, and probed with a 1/2000 dilution of our native wild type polyclonal antibody, as described under "Materials and Methods." *Lane 1*, analysis after oxidative reduction; *lane 2*, after dialysis against the factor Xa buffer; and *lane 3*, after factor Xa cleavage.

DISCUSSION

Currently, the accrual rate of new protein sequences through gene cloning far outstrips the rate of determination of three-dimensional structure. Epo is among a large number of biologically important proteins which have not yet been analyzed by x-ray diffraction or NMR. The problem is simplified by cumulative evidence that the structures of most proteins are likely to be variations on existing themes (27). Indeed, as mentioned above, Epo appears to share common structural features with a large group of cytokines (15–17).

Computer-based prediction of structure can be reduced to a three-stage process: secondary structure is predicted from the primary amino acid sequence and, when available, optical measurements. Analysis of Epo by circular dichroism reveals about 50% α-helix and no detectable β-sheet (7, 11). With the knowledge of disulfide bonds, secondary structural elements are then packed into a set of alternative tertiary structures. The number of plausible arrangements can be reduced by empirical knowledge of preferred helix-helix packing geometries and the need for globular structure to form a hydrophobic core. The putative tertiary structure is then refined by standard force field calculations. Since there are a large number of alternate tertiary structures, the availability of experimentally determined structure of a homologous protein is critically important. Thus, the predicted model of Epo structure gains considerable validity by knowledge of the structures of growth hormone (20, 21) and IL-4 (22, 23).

We have tested the predicted four anti-parallel α-helical bundle structure by means of site-directed mutagenesis. Deletions within predicted α-helices would be expected to destabilize tertiary structure, whereas deletions or insertions in non-helical segments should be permitted unless they impose undue strain on the structure. For example, a deletion in an overhand inter-helical loop may result in insufficient length to connect the two helices. Results that we have obtained on muteins produced in mammalian (Cos7) cells are summarized in Fig. 8. Our measurements of the quantities of processed mutein by RIA may underestimate the true amount of secreted Epo. Even a small deletion or insertion can result in a conformational change that may lead to impaired binding by our polyclonal antibody, raised against native human Epo. Thus the values for specific activity (biologic activity/RIA) that we report must be regarded as approximations. This caveat notwithstanding, our mutagenesis results are in good agreement with our proposed four α-helical model of erythropoietin. The proper folding of Epo into its native tertiary structure is necessary for stability and biological function. Muteins with short deletions inside predicted α-helices were not processed and exhibited no biological activity. In contrast, when deletions were created in predicted interconnecting loops, secreted proteins were detected, to varying degrees, both by radioimmunoassay and bioassay. Furthermore, additions or deletions at the NH_2 or COOH termini did not markedly impair the secretion and the biological activity of the Epo protein. Moreover, mutations at Cys^{29} and Cys^{33} showed that the small disulfide loop is not critical for biological activity. In order to delineate Epo's functionally important residues involved in the direct binding onto the Epo receptor, we have prepared and tested a series of amino acid replacements on the surfaces of the predicted α-helices. These experiments will be described in a subsequent paper.

Acknowledgments—We thank Elizabeth Eldridge for technical assistance and Hanna Bandes for secretarial and editorial assistance. We also thank Dr. David Hankins for providing the HCD57 cell line.

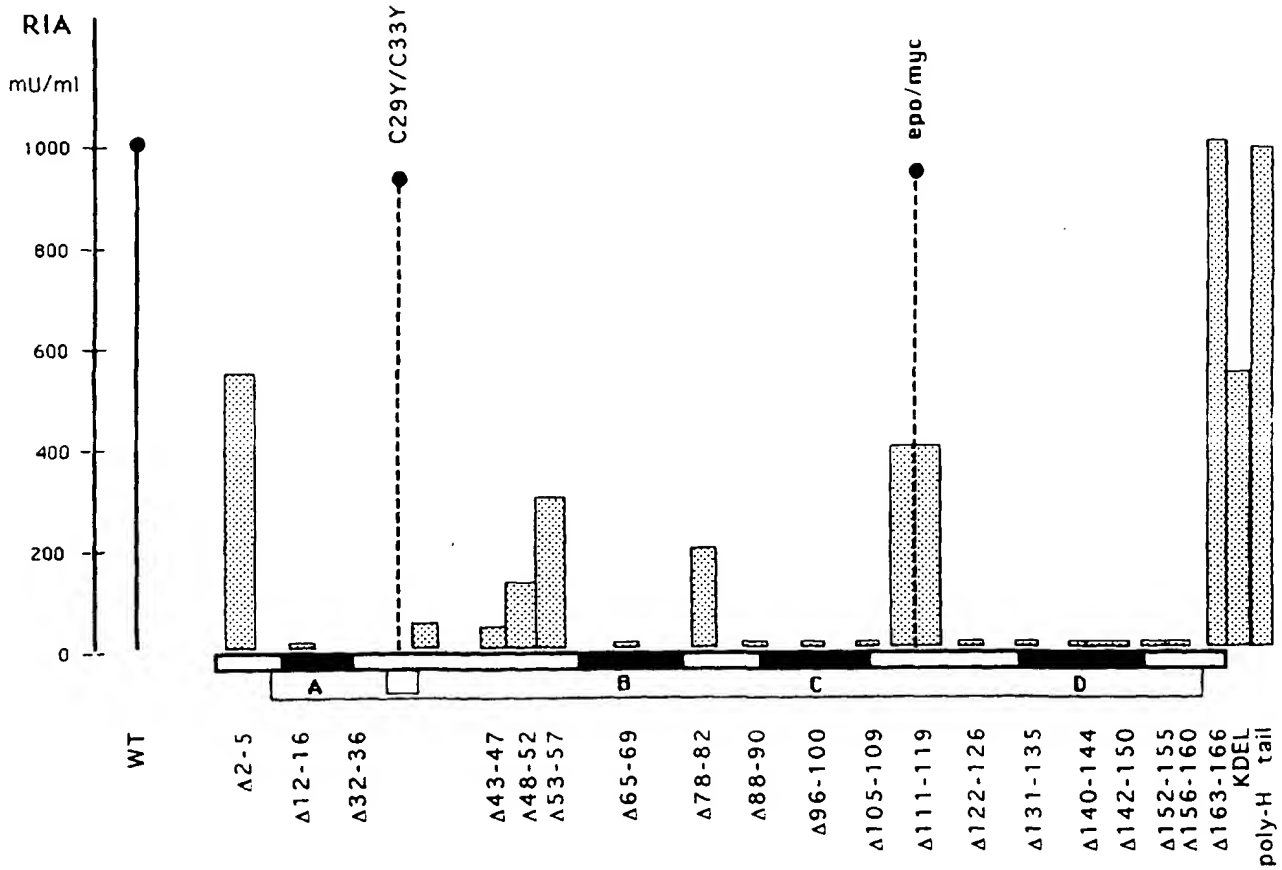


FIG. 8. **Relationship between production of muteins and proposed secondary structure.** This *bar graph* shows the amount of secreted proteins in the supernatants of transiently expressed Epo mutants, as detected by radioimmunoassay. The muteins were aligned over a schematic representation of the native Epo molecule. Each deletion is shown as a *stippled bar*, the width of which is proportional to the number of residues deleted. The four α-helices are represented by the *black rectangles*. The two disulfide bridges are indicated. These mutagenesis results are in good agreement with our proposed four α-helical model of Epo.

Dr. N. Komatsu for the UT-7/Epo cell line, and Dr. John Winkelman for the Epo receptor-agarose beads.

REFERENCES

1. Carnot, P., and Deflandre, C. (1906) *C. R. Seances Acad. Sci. (Paris)* **143**, 432-435
2. Krumdieck, N. (1943) *Proc. Soc. Exp. Biol. Med.* **54**, 14-17
3. Reissmann, K. R. (1950) *Blood* **5**, 372-380
4. Erslev, A. (1953) *Blood* **8**, 349-357
5. Jacobson, L. O., Goldwasser, E., Fried, W., and Plzak, L. (1957) *Nature* **179**, 633-634
6. Miyake, T., Kung, C. K. H., and Goldwasser, E. (1977) *J. Biol. Chem.* **252**, 5558-5564
7. Lai, P.-H., Everett, R., Wang, F.-F., Arakawa, T., and Goldwasser, E. (1986) *J. Biol. Chem.* **261**, 3116-3121
8. Jacobs, K., Shoemaker, C., Rudersdorf, R., Neill, S. D., Kaufman, R. J., Seehva, A., Jones, S. S., Hewick, R., Fritsch, E. F., Kawakita, M., Shimizu, T. and Miyake, T. (1985) *Nature* **313**, 806-810
9. Lin, F.-K., Suggs, S., Lin, C.-H., Browne, J. K., Smalling, R., Egrie, J. C., Chen, K. K., Fox, G. M., Martin, F., Stabinsky, Z., Badrawi, S. M., Lai, P.-H. and Goldwasser, E. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 7580-7584
10. Sasaki, H., Bothner, B., Dell, A., and Fukuda, M. (1987) *J. Biol. Chem.* **262**, 12059-12076
11. Davis, J. M., Arakawa, T., Strickland, T. W., and Yphantis, D. A. (1987) *Biochemistry* **26**, 2633-2638
12. D'Andrea, A. D., Lodish, H., and Wong, G. C. (1989) *Cell* **57**, 277-285
13. Koury, M. J., and Bondurant, M. C. (1990) *Science* **248**, 378-381
14. Bazan, J. F. (1989) *Biochem. Biophys. Res. Commun.* **164**, 788-796
15. Parry, D. A. D., Minasian, E., and Leach, S. J. (1988) *J. Mol. Recognit.* **1**, 107-110
16. Bazan, J. F. (1990) *Immunol. Today* **11**, 350-354
17. Manavalan, P., Swope, D. L., and Withy, R. M. (1992) *J. Protein Chem.* **11**, 321-331
18. Cohen, F. E., and Kuntz, I. D. (1987) *Proteins Struct. Funct. Genet.* **2**, 162-166
19. Curtis, B. M., Presnell, S. R., Srinivasan, S., Sassenfeld, H., Klinke, R., Jeffery, E., Cosman, D., March, C. J., and Cohen, F. E. (1991) *Proteins Struct. Funct. Genet.* **11**, 111-119
20. Abdel-Meguid, S. S., Shieh, H.-S., Smith, W. W., Dayringer, H. E., Violand, B. N., and Bentle, L. A. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 6434-6437
21. de Vos, A. M., Ultsch, M., and Kossiakoff, A. A. (1992) *Science* **255**, 306-312
22. Redfield, C., Smith, L. J., Boyd, J., Lawrence, G. M. P., Edwards, R. G., Smith, R. A. G., and Dobson, C. M. (1991) *Biochemistry* **30**, 11029-11035
23. Powers, R., Garrett, D. S., March, C. J., Frieden, E. A., Gronenborn, A. M., and Clore, G. M. (1992) *Science* **256**, 1673-1677
24. Diederichs, K., Boone, T., and Karplus, A. (1991) *Science* **254**, 1779-1782
25. Pandit, J., Bohm, A., Jancarick, J., Halenbeck, R., Koths, K., and Kim, S.-H. (1992) *Science* **258**, 1358-1362
26. Wen, D., Boissel, J.-P., Tracy, T. E., Gruninger, R. H., Mulcahy, L. S., Czelusniak, J., Goodman, M., and Bunn, H. F. (1993) *Blood*, in press
27. Levitt, M., and Chothia, C. (1976) *Nature* **261**, 552-558
28. Cohen, F. E., Abarbanel, R. M., Kuntz, I. D., and Fletterick, R. J. (1986) *Biochemistry* **25**, 266-275
29. Garnier, J., Osguthorpe, D. J., and Robson, B. (1978) *J. Mol. Biol* **120**, 97-120
30. Chou, P. Y., and Fasman, G. D. (1978) *Annu. Rev. Biochem.* **47**, 251-276
31. Presnell, S. R., Cohen, B. I., and Cohen, F. E. (1992) *Biochemistry* **31**, 983-993
32. Richmond, T. J., and Richards, F. M. (1978) *J. Mol. Biol.* **119**, 537-555
33. Cohen, F. E., Richmond, T. J., and Richards, F. M. (1979) *J. Mol. Biol.* **132**, 275-288
34. Cohen, F. E., Kosen, P. A., Kuntz, I. D., Epstein, L. B., Ciardelli, T. L., and Smith, K. A. (1986) *Science* **234**, 349-352
35. Cohen, F. E., and Kuntz, I. D. (1987) *Proteins Struct. Funct. Genet.* **2**, 162-166
36. Sternberg, M. J. E., and Cohen, F. E. (1982) *Int. J. Biol. Macromol.* **4**, 137-144
37. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., pp. 1.38-1.41, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
38. Kunkel, T. A., Roberts, J. D., and Zakour, R. A. (1987) *Methods Enzymol.* **154**, 367-382
39. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463-5467
40. Kingston, R. E., Chen, C. A., and Okayama, H. (1991) *Current Protocols in Molecular Biology*, Vol. 1, pp. 9.1.1-9.1.9, Greene Publishing Associates and Wiley-Interscience, New York

41. Chirgwin, J. M., Pryzbyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299
42. Faquin, W. C., Schneider, T. J., and Goldberg, M. A. (1992) *Blood*, **79**, 1987–1994
43. Faquin, W. C., Schneider, T. J., and Goldberg, M. A. (1993) *Exp. Hematol.* **21**, in press
44. Laemmli, U. K. (1970) *Nature* **227**, 680–685
45. Krystal, G. (1983) *Exp. Hematol.* **11**, 649–660
46. Goldberg, M. A., Glass, G. A., Cunningham, J. M., and Bunn, H. F. (1987) **84**, 7972–7976
47. Hankins, W. D., Chin, K., Dons, R., and Szabo, J. (1987) *Blood* **70**, 173a
48. Komatsu, N., Nakauchi, H., Miwa, A., Ishihara, T., Eguchi, M., *et al.* (1991) *Cancer Res* **51**, 341–348
49. Yoshimura, A., D'Andrea, A. D., and Lodish, H. F. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 4139–4143
50. Gershoni, J. M., and Palade, G. E. (1983) *Anal. Biochem.* **131**, 1–15
51. Krieg, P. A., and Melton, D. A. (1984) *Nucleic Acids Res.* **12**, 7057–7070
52. Harris, K. W., Mitchell, R. A., and Winkelmann, J. C. (1992) *J. Biol. Chem.* **267**, 15205–15209
53. Presnell, S. R., and Cohen, F. E. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86**, 6592–6596
54. Parry, D. A. D., Minasian, E., and Leach, S. J. (1988) *J. Mol. Recognit.* **1**, 107–110
55. Narhi, L. O., Arakawa, T., Aoki, K. H., Elmore, R., Rohde, M. F., Boone, T., and Strickland, T. W. (1991) *J. Biol. Chem.* **266**, 23022–23026
56. Cosman, D., Lyman, S. D., Iolzerda, R. L., Beckman, M. P., Park, L. S. Goodwin, R. G., and March, C. J. (1990) *Trends Biochem. Sci.* **15**, 265–270
57. Bazan, J. F. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 6934–6938
58. Bowie, J. U., Luthy, R., and Eisenberg, D. (1991) *Science* **253**, 164–170
59. Smith, L. J., Redfield, C., Boyd, J., Lawrence, G. M., Edwards, R. G., Smith, R. A., and Dobson, C. M. (1992) *J. Mol. Biol.* **224**, 899–904
60. Abdel-Meguid, S. S., Shieh, H.-S., Smith, W. W., Dayringer, H. E., Violand, B. N., and Bentle, L. A. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 6434–6437
61. Diederichs, K., Boone, T., and Karplus, P. A. (1991) *Science* **254**, 1779–1782
62. Emini, E. A., Hughes, J. V., Perlow, D. S., and Boger, J. (1985) *J. Virol.* **55**, 836–839
63. DeLorme, E., Lorenzini, T., Griffin, J., Martin, F., Jacobsen, F., Boone, T., and Elliott, S. (1992) *Biochemistry* **331**, 9871–9876
64. Wasley, L. C., Timony, G., Murtha, P., Stoudemire, J., Dorner, A. J., Caro, J., Krieger, M., and Kaufman, R. J. (1991) *Blood* **77**, 2624–2632
65. Andres, D. A., Rhodes, J. D., Meisel, R. L., and Dixon, J. E. (1991) *J. Biol. Chem.* **266**, 14288–14292
66. Wang, F. F., Kung, C. K.-H., and Goldwasser, E. (1985) *Endocrinology* **116**, 2286–2292
67. Lin, F.-K. (1987) *Molecular and Cellular Aspects of Erythropoietin and Erythropoiesis: NATO ASI Series*, Vol. H8, pp. 23–36, Spinger-Verlag, Berlin
68. Yamaguchi, K., Akai, K., Kawanashi, G., Veda, M., Masuda, S., and Sasaki, R. (1991) *J. Biol. Chem.* **266**, 20434–20439
69. Dubé, S., Fisher, J. W., and Powell, J. S. (1988) *J. Biol. Chem.* **263**, 17516–17526
70. Takeuchi, M., Takasaki, S., Shimada, M., and Kobata, A. (1990) *J. Biol. Chem.* **265**, 12127–12130
71. Fukuda, M. N., Sasaki, H., Lopez, L., and Mukuda, M. (1989) *Blood* 73, 84–89
72. Tsuda, E., Kawanishi, G., Veda, M., Masuda, S., and Sasaki, R. (1990) *Eur. J. Biochem.* **188**, 405–411
73. Sasaki, H., Bothner, B., Dell, A., and Fukuda, M. (1987) *J. Biol. Chem.* **262**, 12059–12076
74. Spivak, J. L., and Kogans, B. B. (1989) *Blood* **73**, 90–99

Fine-Structure Epitope Mapping of Antierythropoietin Monoclonal Antibodies Reveals a Model of Recombinant Human Erythropoietin Structure

By Steve Elliott, Tony Lorenzini, David Chang, Jack Barzilay, Evelyne Delorme, James Giffin, and Lyndal Hesterberg

We have isolated and mapped the rHuEPO epitopes for three noncompeting anti-EPO monoclonal antibodies (MoAbs). The MoAb 9G8A recognizes a linear epitope that includes amino acids 13, 16, and 17. MoAb F12 recognizes a conformational epitope that includes amino acids 31 through 33, 86 through 91, and 138. MoAb D11 recognizes a conformational epitope that includes amino acids 64 through 78 and 99 through 110. MoAb D11 neutralizes rHuEPO activity which suggests that its epitope may contain the receptor binding domain. Analysis of the effect of mutations on folding allowed the identification of buried residues, α-helical, and non α-helical regions. This data along with epitope mapping data of anti rHuEPO monoclonals was used to model rHuEPO protein structure. A model consistent with the data is a 4-helix bundle with short and long interconnecting loops.

THE ERYTHROPOIETIN (EPO) gene has been cloned and biologically active material (rHuEPO) has been produced in mammalian cells.[1-4] The gene encodes a 193 amino acid protein.[1,2] The secreted protein is 165 amino acids with approximately 40% carbohydrate.[3,5] Several models of EPO structure have been described[6-9] but there is no available nuclear magnetic resonance (NMR) or crystal structure. One model[7] is supported by analysis of the effect of deletions on secretion of EPO. In this case it is assumed that deletion of helices, but not connecting loops, will result in a decreased rate of secretion because of a loss of EPO structure. Unfortunately there is little additional information to test existing models or to develop new ones. Therefore we decided to develop an immunological approach to determine rHuEPO structure.

Immunological studies can be used to determine structural and functional information about proteins. For example, mapping of epitopes recognized by neutralizing antibodies reveals information about the active site. Conformational epitopes require proper folding to juxtapose the amino acids involved in antibody binding. Therefore, epitope mapping by antibodies that recognize conformational epitopes reveals structural information about the antigen because amino acid residues close together in space are identified. When epitope mapping by several different antibodies is done it is possible to determine the folding pattern of the molecule. Additional constraints on structure can be obtained if the antibodies recognize nonoverlapping epitopes. The epitopes must therefore be separated from each other in the structural model. This type of approach has been used successfully to develop a model of human growth hormone structure.[10]

There are many different approaches to mapping of epitopes. Antibody binding to synthetic peptides or peptide fragments from the antigen can localize the binding site. This approach is usually successful only for antibodies that recognize linear epitopes. Another approach is to determine antibody binding to naturally occurring variants, such as proteins from different species. However, such variants may not be available or the variants may have too many amino acid differences to unambiguously identify those important for antibody binding. In addition, this approach may also fail if the antibody recognizes a particular conformation and the variants have different conformations. A third approach is to construct variants with single amino acid substitutions by recombinant DNA methodologies and then determine which mutations result in reduced antibody binding. This approach can result in false positives if the mutation alters folding in a way that affects antibody binding. However, it has the advantage that mutations in residues that are part of the epitope result in reduced binding even for antibodies that recognize conformational epitopes. Thus, if one can distinguish between specific and nonspecific effects on antibody binding, the epitope can be determined.

Our approach to mapping antibody epitopes was to construct rHuEPO variants and measure antibody binding to them. We used two methods to overcome the possibility that mutations would result in false positives because of changes in rHuEPO conformation that indirectly affected immunoreactivity. In many cases multiple substitutions were made at each position to increase the probability that a mutation could be identified that had a minimal effect on protein folding. Secondly, we also attempted to identify those mutations that resulted in altered protein folding. This was done by assaying the rHuEPO variants with conformation sensitive antibodies including the anti-EPO monoclonal antibody (MoAb), 9G8A. 9G8A immunoreactivity increases when rHuEPO is denatured.[11] Thus, mutations that reduce MoAb immunoreactivity because they cause gross conformational changes were identified. Another benefit of this approach is that predictions of secondary structure can be made according to the effect of mutations on folding. For example, it has been reported that mutations in connecting loops have little effect on structure while mutations in α-helical or β-sheet regions alter structure.[12-14] We have developed a model of rHuEPO structure by using the epitope mapping data in combination with predictions according to effects of amino acid substitutions on folding.

MATERIALS AND METHODS

Materials. EPO peptides were synthesized on an Applied Biosystems (Foster City, CA) model 430A peptide synthesizer. The

From Amgen Inc, Thousand Oaks, CA.

Submitted March 31, 1995; accepted November 9, 1995.

Address reprint requests to Steve Elliott, PhD, Amgen Inc, Amgen Center, Thousand Oaks, CA 91320.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. section 1734 solely to indicate this fact.

0006-4971/96/8707-0048$3.00/0

isolation and characterization of the MoAbs F12, D11, and 9G8A has been described.[11] The EPO standard used in immunoassays is recombinant human EPO produced in Chinese Hamster Ovary cells. ^{125}I EPO was obtained from Amersham (Arlington Heights, IL).

Radioimmunoassay (RIA) and enzyme-linked immunosorbent assay (ELISA). RIA-P uses an anti-EPO polyclonal antibody raised in rabbits and has been described previously.[4] RIA-N[15] uses a rabbit polyclonal antibody raised against a synthetic peptide (amino acids 1 through 20 of human EPO). This assay appears to be largely independent of folding of rHuEPO. The ELISA assays used immobilized MoAbs (F12 or D11) as capture antibodies. A secondary antibody raised in rabbits (R) or Goats (G) was labeled with horse radish peroxidase and was used as a signal antibody. RIA assays with 9G8A and ELISA assays with F12 (EIA/F12) and D11 (EIA/D11) are described in an accompanying report.[11]

Epitope mapping of 9G8A with peptides. Synthetic EPO peptides, in 0.05 mol/L carbonate/bicarbonate pH 9.2 buffer, were bound to polystyrene wells (Costar, Cambridge, MA) at room temperature (RT) for 2 hours. The plate was incubated at RT for 2 hours then overnight at 4°C. The next day the wells were blocked for 30 minutes at RT with phosphate-buffered saline (PBS) (10 mmol/L Na · PO4/ 150 mmol/L NaCl pH 6.8) containing 5% bovine serum albumin (BSA) followed by a 2-hour RT incubation with protein A agarose purified 9G8A at a concentration of 5 ug/mL in PBS containing 1% BSA. After washing with KPL washing buffer (Kirkegard and Perry Labs, Inc Gaithersburg, MD) the plate was incubated with a 1:1,000 dilution of goat-antimouse IgG conjugated with horse radish peroxidase (Boehringer Mannheim, Indianapolis, IN) for an hour at RT. The plate was then washed and developed with ABTS (2,2'-azino-di-[3-ethylbenzthiazoline sulfonate]) (Kirkegard and Perry Labs, Inc) substrate solution. Colorimetry was conducted at 405 nmol/L.

Competition assays with synthetic peptides used for 9G8A epitope mapping were performed by immobilizing 50 ng EPO in carbonate buffer for 2 hours at RT to polystyrene wells. The contents were then dumped and the wells were blocked for 1 hour with PBS/5% BSA. After removal of the contents, 50 μL peptide and 50 μL 9G8A (100 ng/mL) were simultaneously added. Incubation was for 3 hours at 4°C. The wells were then washed with KPL wash buffer, incubated with ^{125}I labeled goat-antimouse antibodies, washed again, and counted.

Construction of EPO variants. All rHuEPO variants were expressed from an human EPO cDNA clone[16] in an SV-40 vector.[1] To construct rHuEPO variants, a DNA segment containing EPO coding sequences was transferred to M13mp18[17] and single stranded DNA was produced in a *dut ung Escherichia coli* strain (RZ1032) as described.[18,19] The single-stranded DNA was annealed with synthetic oligonucleotides containing the desired mutations and then was extended with Klenow polymerase in the presence of nucleotide triphosphates and T4 DNA ligase. After an overnight incubation, the reaction mixture was transformed into wild type *E coli* (JM109;[17]). Transfectants were screened with the mutant oligonucleotides to identify mutants. The presence of each mutation was confirmed by DNA sequence analysis then the DNA segment containing the desired mutation was transferred back to the original SV-40 vector.

Transfection of COS-1 cells. COS-1 cells were transfected by one of two methods. The calcium phosphate method has been described previously.[20] To transform by electroporation, cells were procured from semiconfluent dishes, washed in Dulbecco's modified Eagle's medium (5% fetal calf serum, 1% L-glutamine/penicillin/ streptomycin, Irvine Scientific, Irvine, CA) and resuspended at 4×10^6 cells per mL. One mL of cells was transferred to an electroporation cuvette, and electroporated at 25 μfarads and 1,600 V in the presence of 2 through 20 μg plasmid DNA. The electroporated cells were then plated at 2×10^6 cells per 60mm tissue culture dish in 5 mL of medium. Two to 4 hours later the medium was removed and replaced with 5 mL fresh medium. With both transfection methods the conditioned medium (approximately 5 mL) was collected after 3 to 5 days. Aliquots were made and stored at −70°C.

RESULTS

Construction of rHuEPO variants for epitope mapping. To identify amino acids involved in antibody binding we constructed 207 rHuEPO variants with individual point mutations that altered 131 of the 165 amino acid positions in rHuEPO as well as 2 positions in the signal peptide (Figs 1 and 2). The largest space between mutations was 3 amino acids. At 56 amino acid positions, 2 or more variants with different substitutions were made. In general, the mutations were nonconservative substitutions. This was done to ensure that if a region was important for antibody binding or rHuEPO structure, the mutation would result in an effect on antibody immunoreactivity. In cases where the nonconservative substitution was thought to affect protein folding, additional conservative substitutions were also made. With this approach we hoped to identify structural elements in rHuEPO by distinguishing between regions that could and could not tolerate nonconservative substitutions and still have confidence that we could identify antibody epitopes.

Expression of rHuEPO variants and determination of concentration. Because of the large number of rHuEPO variants needed for epitope mapping and the desire to increase the probability that proteins would be properly folded, we decided to express the variants transiently in COS cells and collect and assay conditioned medium containing secreted rHuEPO. In many cases several independent transfections for each rHuEPO variant were done. Because of the low expression levels in this system as well as the large numbers of variants, it was not practical to purify each variant. We chose instead to determine the concentration of rHuEPO in each sample directly by immunoassay.

The concentration of each rHuEPO variant in supernatants was determined by five different immunoassays including RIA-P, RIA-N, EIA/D11, EIA/F12, and RIA-9G8A immunoassays. ELISA assays were performed using D11 and F12 as primary MoAbs in combination with rabbit (R) and, in some cases, goat (G) horseradish peroxidase conjugated signal polyclonal antibodies.[11] The use of two different secondary antibodies raised in different species reduced the likelihood that low immunoreactivity in ELISAs was because of reduced immunoreactivity by the secondary antibody. The redundancy in assays allowed increased confidence that at least one assay could be found that was largely unaffected by the mutation. The fact that 9G8A, F12, and D11 all recognized nonoverlapping epitopes suggested that a rHuEPO variant with a mutation that affected the epitope for one antibody would not reduce the immunoreactivity of the other immunoassays. This would be especially true for mutations having little or no effect on rHuEPO conformation. In addition two different RIA assays using polyclonals were also performed. RIA-P uses a polyclonal antibody raised to rHuEPO, thus most rHuEPO mutations would be unlikely to directly affect immunoreactivity with this assay. RIA-N

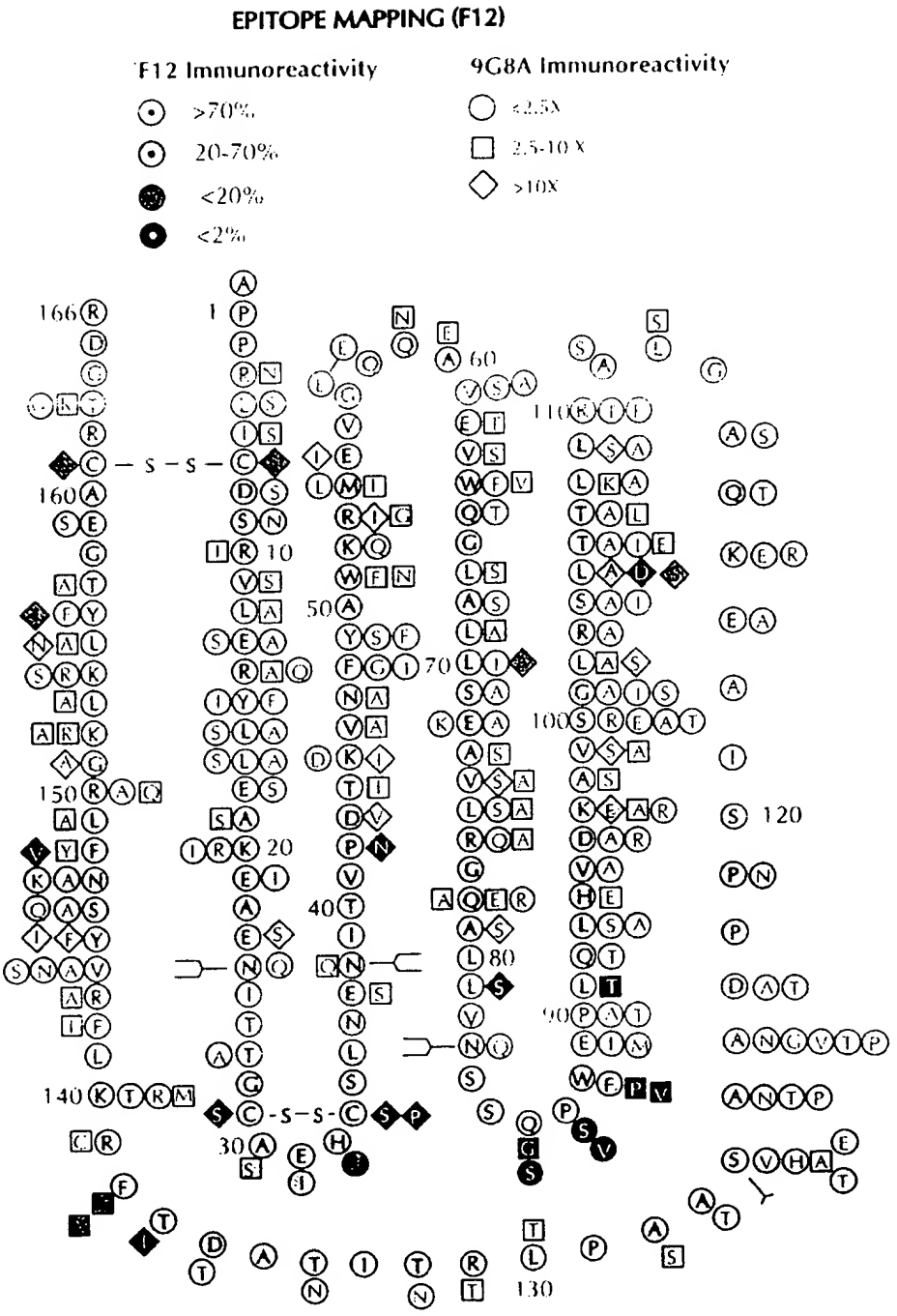


Fig 1. Fine structure epitope mapping of anti-EPO MoAb F12. Amino acid residues in the human EPO sequence (gray letters) and EPO variants (adjacent to human sequences [black letters]) are indicated. Mutations connected by a line indicate that both were made simultaneously. Branched symbols at amino acids 24, 38, 83, and 126 represent the locations of glycosylation sites. The shape of the symbols surrounding the mutations indicate the effect on 9G8A immunoreactivity. Shaded symbols indicate mutations that result in decreased F12 immunoreactivity. Immunoreactivity was determined as described in the text.

used a polyclonal antibody raised to an amino terminal (aa 1-20) HuEPO peptide. This assay is largely unaffected by rHuEPO conformation and, in most cases, estimates of EPO concentrations determined by this assay are likely to reflect true concentrations.

It is likely that many mutations in rHuEPO would alter the folding. Furthermore the fact that both F12 and D11 recognize conformational epitopes suggests that some rHuEPO mutations would indirectly affect the immunoreactivity to these monoclonals. To identify such mutations, several analyses were performed. First, the concentration of rHuEPO in supernatant was measured by an assay thought to be unaffected by the mutation to see if the secretion level might be lower than expected. Reduced secretion is thought to be indicative of altered or incomplete folding. In numerous experiments with native sequence rHuEPO we found that

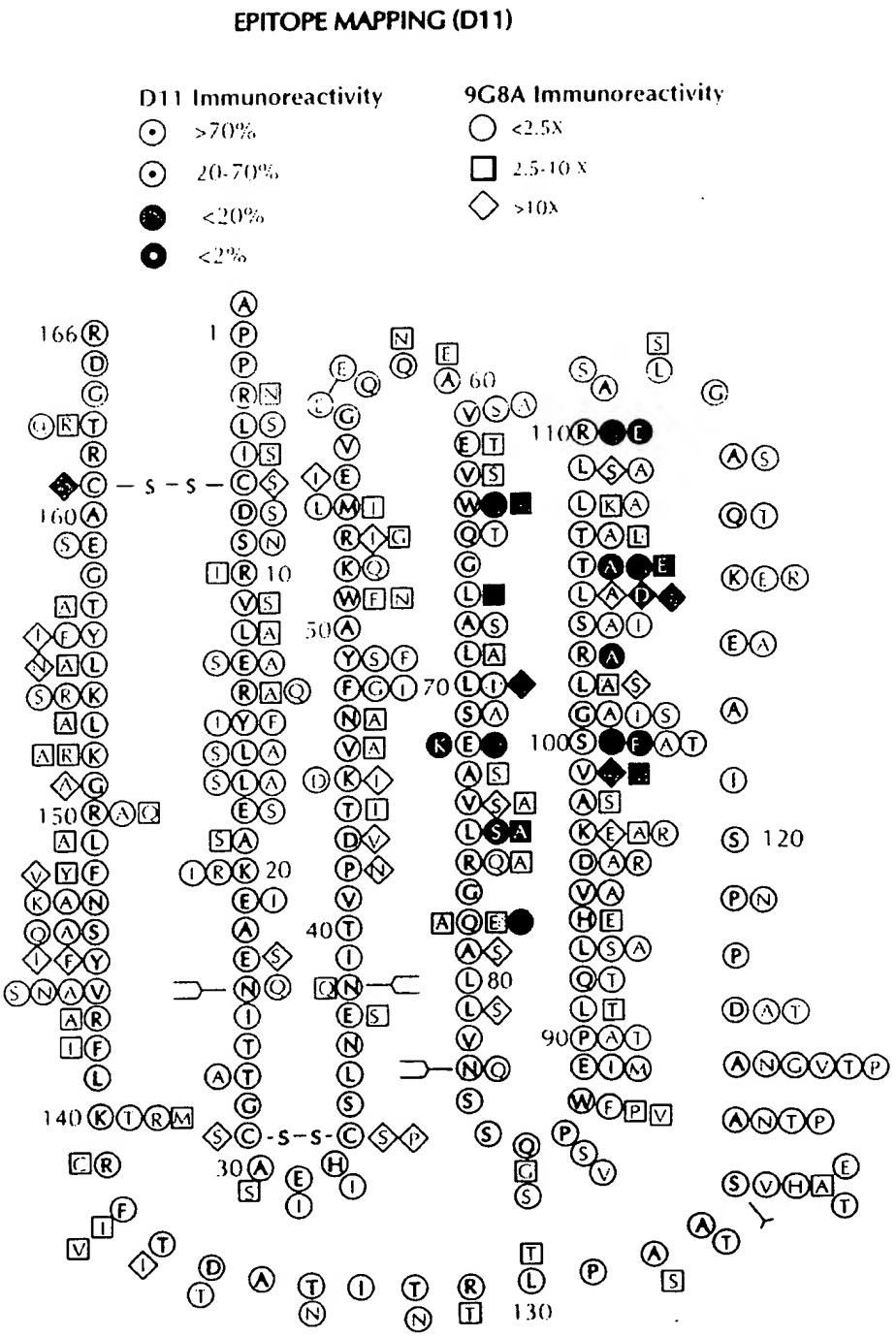


Fig 2. Fine structure epitope mapping of anti-EPO MoAb D11. Amino acid residues in the human EPO sequence (gray letters) and EPO variants (adjacent to human sequences [black letters]) are indicated. Mutations connected by a line indicate that both were made simultaneously. Branched symbols at amino acids 24, 38, 83, and 126 represent the locations of glycosylation sites. The shape of the symbols surrounding the mutations indicate the effect on 9G8A immunoreactivity (fold increase relative to RIA-N and or other immunoassays). Shaded symbols indicate mutations that result in decreased D11 immunoreactivity. Immunoreactivity was determined as described in the text.

conditioned medium from COS cells contained 10 to 100 U/mL rHuEPO. Therefore, levels less than 5 U/mL were thought to be indicative of altered folding and thus reduced secretion. Secondly, immunoassays thought to be affected by changes in rHuEPO conformation were examined. RIA-P can have reduced apparent immunoreactivity whereas RIA-9G8A immunoreactivity has increased immunoreactivity when rHuEPO conformation is altered.[11] In cases where RIA-P was low and 9G8A was high, EIA/F12-R and EIA/D11-R were also often low. In these cases RIA-N immunoreactivity was usually much higher than with any other immunoassay. Thus, we concluded that rHuEPO conformation

Table 1.

Mutation	WT AA	Immunoreactivity RIA (U/mL)† P	N	9G8A	EIA-F12‡ (U/mL) R	G	EIA-D11§ (U/mL) R	G
wt	wt	33	26	31	35	35	36	39
wt	wt	66	60	53	73	60	70	71
pro-2*	leu	29	<0.05	57	40		26	
asn4	arg	3.6	2.3	9	3.8	3.4	3.4	2.9
ser5	leu	6.2	<1.1	8.5	10.7		12.2	
ser7	cys	0.6	12	263	0.2	0.6	0.4	11
ala12	leu	1.8	2.8	12	1.7		1.7	2.9
ala13	glu	4.4	1.9	2.3	4.7	3.4	3.8	0.9
ser13	glu	12.6	14.8	<1	14.2		11.1	
gln14	arg	19	20	17	20		20	
ile15	tyr	10	12	8	11		8.3	12
ser16	leu	11	12.9	1.5	12.2		13	
ser17	leu	8.8	11.7	3.8	12.7		12.4	
ser29	cys	1.4	2.8	39	<0.07	<0.06	1.7	3.0
ile32	his	10	10	10	0.6		11	10
pro33	cys	2.3	12	280	<0.2	<0.5	2.8	15
ser70	leu	<0.14	0.38	4	<0.06		<0.06	
lys72	glu	15	10	11	17	17	<0.15	
ser79	ala	3.7	2.8	34	2.9		2.1	2.6
ser81	leu	9.6	22	290	<0.06		10.7	10.3
gly86	gln	47	34	172	<0.06		51	37
glu100	ser	82	63	125	84	84	0.7	<1
ala103	arg	7.9	6.9	7	8.2		<0.06	
ile104	ser	21	18	56	24		23	
ser105	leu	0.17	1.9	32	0.1		0.2	<0.3
glu106	thr	28	31	108	27		<0.06	
ile138	phe	5.6	3.4	76	0.19		4.8	7.1
val138	phe	1.6	1.7	7.0	<0.07		0.9	1.3
phe145	tyr	25	25	305	23		29	38
val148	phe	0.3	2.5	69	<0.06		0.3	
ala151	gly	5	24	203	6		4	8.6
ile156	tyr	6.5	28	468	7.7	10	19	20

* Number represents amino acid position in rHuEPO. Pro^{-2} is a mutation in the signal peptide.

† RIA was performed as described in Materials and Methods using either (N) N-terminal antibody (P) polyclonal antibody or (9G8A) MoAb.

‡ ELISA with MoAb F12 were performed as described in Materials and Methods using either (R) Rabbit polyclonal secondary antibodies or (G) Goat polyclonal secondary antibodies.

§ ELISA with MoAb D11 were performed as described in Materials and Methods using either (R) Rabbit polyclonal secondary antibodies or (G) Goat polyclonal secondary antibodies.

was altered. RIA-9G8A/total ratios of 0.7 to 2.5 were found with samples containing folded rHuEPO and ratios of 20 through 40 correlated with completely denatured rHuEPO.

Total antigen concentration was calculated for each rHuEPO variant and was the average of all the assays that were unaffected by the mutations. Ratios were then calculated for each immunoassay; immunoreactivity measured with each assay divided by the total antigen concentration. As an example of this analysis a Cys^{33} to pro substitution resulted in secretion of 12 U/mL as measured by RIA-N and 15 U/mL as measured by EIA/D11-G (Table 1). Therefore we assume the rHuEPO concentration was 12 to 15 U/mL. This concentration is reasonable because it was in the normal range. RIA-P immunoreactivity was 5× lower and 9G8A immunoreactivity was 23× higher suggesting that rHuEPO conformation is altered with this mutation. This is not unexpected because Cys^{33} is part of a disulfide bond that may be important for rHuEPO structure. F12 immunoreactivity was low suggesting that this assay is also affected by the mutation. Immunoreactivity measured by EIA/D11-R was lower than that of EIA/D11-G. This suggests that monoclonal D11 and the goat secondary antibody were capable of binding to this rHuEPO variant but the immunoreactivity measured by the rabbit secondary antibody was reduced. Using this type of analysis, we could determine both the concentration of each variant in samples and the degree to which each mutation affected immunoreactivity and the conformation of rHuEPO. Reductions of 30% or more in apparent F12 or D11 immunoreactivity were considered significant. Mutations thought to change rHuEPO conformation were defined as those that increased 9G8A immunoreactivity by more than 2.5 times expected values.

Analysis of RIA-N immunoreactivity. The antibody used in RIA-N was raised against a peptide containing the first 20 amino acids of HuEPO.[15] It appeared to react normally to all the rHuEPO variants with mutations between 21 and 166 because immunoreactivity was always equal to or higher than immunoreactivities measured with all other immunoassays except RIA-9G8A. Although the antibody in RIA-N is a polyclonal it is possible that rHuEPO variants with mutations in the amino terminal 20 amino acids may affect RIA-N immunoreactivity. Mutations at amino acid residues 6, 7, 8, 10, 11, 12, 14, 15, 16, 17, 18, and 20 all gave equal or higher EPO concentrations measured with RIA-N, to that obtained with all other immunoassays except RIA-9G8A. However, compared to some other immunoassays a Ser^5 substitution reduced immunoreactivity over 10-fold and Asn^4 and Asn^9 mutations reduced immunoreactivity approximately 30 to 35%. An Ala^{13} substitution reduced RIA-N immunoreactivity approximately twofold whereas a Ser^{13} mutation had no apparent effect. It has been reported elsewhere that immunoreactivity to monkey EPO with RIA-N is reduced 150-fold.[11] Monkey EPO is extended by 3 amino acids (V-P-G) because of altered processing of the signal peptide,[21] presumably because of the pro residue at −2. Monkey EPO differs at one other position in the first 20 amino acid residues, position 16. However, a rHuEPO Leu^{16} to Ser variant appears to be recognized normally by RIA-N (Table 1). To determine whether alterations at the amino terminus explain the reduced immunoreactivity to monkey EPO, a rHuEPO variant with a pro substitution at −2 was constructed and assayed. Immunoreactivity with RIA-N was reduced over 500-fold with this variant (Table 1). The apparent immunoreactivity using other immunoassays suggested that this mutant was secreted normally (26 to 40 U/mL). In addition, the immunoreactivity with 9G8A was also in this range 59 U/mL. These results suggest that this protein is secreted and folded normally. A different substitution, Leu to Val at −11, in the signal peptide had no apparent effect on immunoreactivity with RIA-N (data not shown). These results suggest that processing of the signal peptide may be altered by a pro substitution at −2 and this alteration affects RIA-N immunoreactivity. Thus, the only mutations that dramatically affected RIA-N (>10-fold decrease) were at the amino terminus and position 5. Residues C-terminal to position 5 appeared to affect RIA-N modestly (twofold or less). This suggests that the major epitope for RIA-N is contained within the first 5 residues.

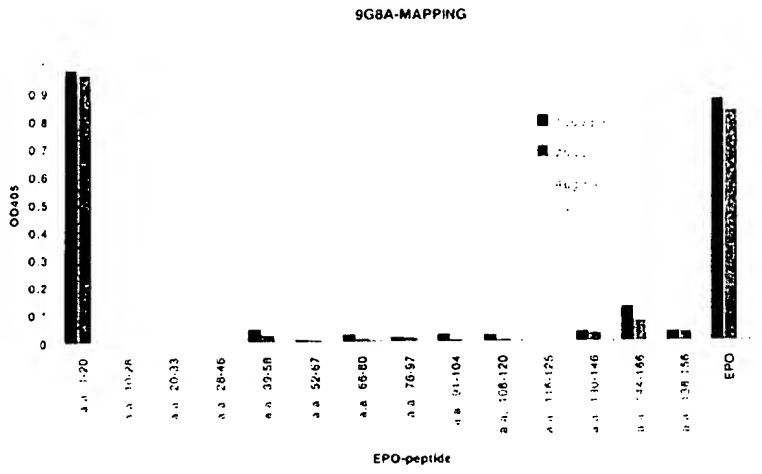


Fig 3. Epitope mapping of anti-EPO MoAb 9G8A with peptides. The indicated peptides at the concentrations shown were bound to plates then incubated with 9G8A and ^{125}I labeled antimouse IgG. Bound ^{125}I was determined as described in Materials and Methods.

Mapping of the 9G8A epitope. The observation that 9G8A can recognize EPO on Western blots and would react with denatured EPO in RIAs suggested that 9G8A recognized a linear epitope. Therefore, we measured 9G8A binding to synthetic peptides with the hope that we could determine the approximate location of the epitope. A collection of overlapping synthetic EPO peptides were bound to microtiter plates and then 9G8A was added to see if any specific antibody binding could be observed as described in Materials and Methods. One peptide that included amino acids 1 through 20 gave a clear positive signal (Fig 3). In this experiment the maximum signal was obtained even with the lowest amount of peptide. All the other peptides resulted in low levels of binding including a peptide containing residues 10 through 28. To further substantiate the role of the amino terminal region in 9G8A immunoreactivity, the ability of the peptides to inhibit binding of rHuEPO in RIA-9G8A was measured. As shown in Fig 4 only the peptide containing amino acid residues 1 through 20 would inhibit rHuEPO binding and it did so in a dose dependent manner. These results suggest that residues between 1 and 20 contain the 9G8A epitope.

To localize the 9G8A epitope further the 205 rHuEPO variants shown in Figs 1 and 2 were examined to see which had reduced 9G8A immunoreactivity. As expected no mutations in residues between 21 and 166 reduced 9G8A immunoreactivity relative to other assays. However, 3 positions in the 1 through 20 region appeared to be important for immunoreactivity including Glu^{13}, Leu^{16}, and Leu^{17} (Table 1, Fig 5). The Ala^{13} mutation resulted in a twofold decrease, Ser^{13} to over 10-fold decrease, Ala^{16} to a 10-fold, Ser^{16} to an eightfold, and Ala^{17} and Ser^{17} to threefold decreases in apparent 9G8A immunoreactivity. Mouse EPO differs from human EPO only at position 16 in the first 20 residues and 9G8A immunoreactivity to mouse EPO is reduced approximately 500-fold.[11] This is consistent with the suggestion that the major epitope for 9G8A includes residues 13 to 17.

Fine structure epitope mapping of MoAb F12. Results from the accompanying report[11] indicate that MoAbs D11 and F12 recognize conformational epitopes. Thus, attempts to map the epitopes with synthetic peptides were unsuccessful. We chose instead to determine the antibody immunoreactivity to closely related rHuEPO variants using the strategy previously described. Mutations at 21 of the 131 amino acid

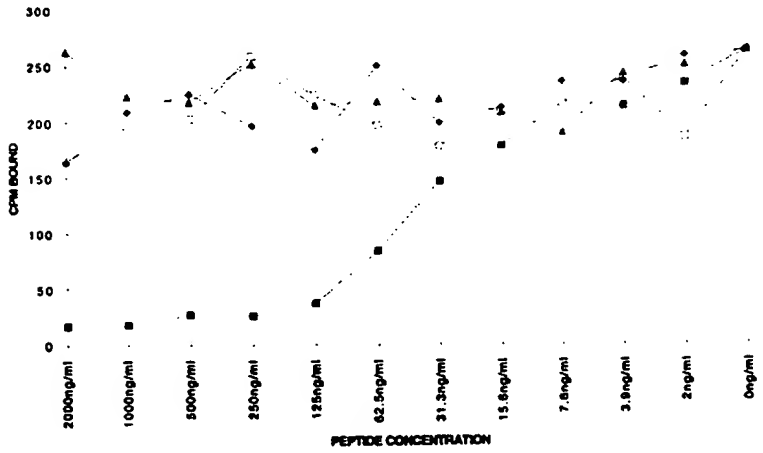


Fig 4. Inhibition of rHuEPO binding in 9G8A immunoassay by synthetic HuEPO peptides. RIA-9G8A was performed by immobilizing rHuEPO to a microtiter plate. The plates were then incubated with the indicated amounts of EPO peptides and 9G8A. The relative amount of 9G8A subsequently bound (CPM) to plates was determined by a final incubation with ^{125}I labeled goat antimouse IgG as described in Materials and Methods. EPO Peptides were as 1 through 20 (■), 41 through 57 (□), 56 through 80 (♦), and 144 through 166 (▲).

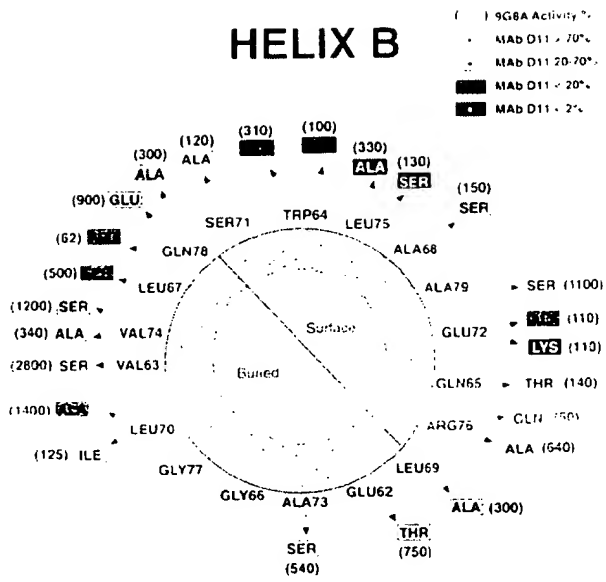


Fig 5. Helical wheel presentation of amino acids 62 to 79 (HELIX B). Amino acid residues (closest to the circle) represent those found in the human EPO sequence. Arrows point towards mutations (in boxes). Mutations that reduced D11 immunoreactivity are shaded as indicated. The effect of each mutation on 9G8A immunoreactivity is shown in parentheses. 9G8A immunoreactivity values are shown as percent of total antigen present as determined by immunoassay (see text). The predicted buried and exposed surfaces of the helix are shown.

positions in rHuEPO tested resulted in reduced F12 immunoreactivity. Examples that show reduced F12 immunoreactivity are shown in Table 1. These and other mutations that caused reduced F12 immunoreactivity are indicated in Fig 1. The mutations that affected F12 immunoreactivity mapped to amino acid residues 7, 29 to 33, 42, 70, 81, 86 to 89, 91, 97, 105, 137, 138, 148, and 156 and 161.

Several of the mapped positions are unlikely to be part of the F12 epitope. Residues 7 and 161 make up a disulfide bond important for structure and activity.[11] The mutations at Pro^{42}, Leu^{70}, Leu^{91}, Lys^{97}, Thr^{137} and Phe^{148} and Tyr^{156} resulted in high 9G8A immunoreactivity, low RIA-P immunoreactivity and, for most, were secreted poorly. This suggests that mutations in these residues affected protein folding and therefore these positions may not be part of the F12 epitope. Variants with mutations at Leu^{81} and Phe^{138} also had high 9G8A immunoreactivity suggesting that rHuEPO structure was affected by the mutations. However, Val^{138} and Ile^{138} variants were recognized similarly by RIA-N, RIA-P, EIA/D11, and the Ser^{81} variant was secreted at normal levels. Thus mutations in these positions affect rHuEPO structure somewhat. However, immunoreactivity was reduced 30-fold or more with F12 which suggests that residues 81 and 138 are close to or part of the epitope. In support of this possibility is the observation that other mutations affected folding (over 10-fold increases in 9G8A immunoreactivity) but they did not affect F12 immunoreactivity eg, Ser^{23}, Val^{43}, Ile^{55} and Ile^{145}. Mutations in the other two regions with low F12 immunoreactivity, amino acids 31, 32, and 87, 88, had normal RIA-P, D11, and 9G8A immunoreactivity. Thus, these residues appear to be directly involved in F12 binding. Residues 29 and 33 make up a disulfide bond. Mutations in these residues result in reduced F12 immunoreactivity but also have low RIA-P and increased 9G8A immunoreactivity suggesting an effect on folding. Given the closeness to residues thought to interact directly with F12 (31 and 32) we cannot rule out their interaction with F12 as well.

Fine structure epitope mapping of D11. The EPO variants were also tested as previously described for D11 immunoreactivity to identify its epitope. Mutations at 21 of the 131 amino acid positions tested resulted in decreased apparent D11 immunoreactivity (Table 1 and Fig 2). These were localized to two main regions: amino acids 64 through 81 and 96 through 110. A Ser mutation at Cys^{161} and a Thr mutation at Leu^{91} also affected D11 immunoreactivity.[11] However 9G8A immunoreactivity was high and RIA-P was low. Cys^{161} forms a disulfide bond with Cys^{7}. Thus, mutations at both positions should result in decreased immunoreactivity with D11 if the decrease was because of disruption of the disulfide bond. However, a Cys^{7} to Ser mutation appears not to affect D11-G (goat secondary antibody) immunoreactivity but D11-G immunoreactivity is reduced at least 20-fold with Cys^{161} mutations.[11] This suggests that Cys^{161} may be close to or part of the D11 epitope whereas Cys^{7} is not. Alternatively, the Ser^{161} mutation may be more destabilizing than the Ser^{7} mutation. In support of the latter possibility is the observation that the Ser^{161} variant is consistently secreted at much lower levels than the Ser^{7} variant.[11]

Identification of amino acids that may be part of a hydrophobic core. Certain mutations in rHuEPO appeared to severely affect structure as evidenced by three criteria: poor secretion, low RIA-P immunoreactivity, and 10-fold or greater than expected 9G8A immunoreactivity. These mutations also resulted in low in vitro bioactivity (data not shown). The residues with this property include Asn^{42}, Leu^{67}, Leu^{70}, Leu^{105}, Phe^{148}, Lys^{152} and Tyr^{156}. These positions contained mostly hydrophobic residues and substitutions of polar residues had much greater effects than did substitutions with hydrophobic residues. For example, a Leu^{70} to Ile mutation did not affect RIA-P, 9G8A, or secretion levels but mutations to Ala or Ser did (Figs 1, 2, and 5). We conclude that these positions may be important for overall structure or be part of the hydrophobic core.

Identification of regions containing α-helices and connecting loops. It has been suggested that mutations in α-helices alter secondary and tertiary structure while mutations in connecting loops are tolerated.[12-14] We hypothesized that we may be able to identify such regions according to effects of the mutations on rHuEPO folding as measured by 9G8A immunoassay. Most regions contained mutations that affected 9G8A immunoreactivity (see Figs 1 or 2). However, there was one stretch of amino acids, 114 through 134, where mutations appeared to have little effect on 9G8A immunoreactivity. Thus, residues 114 through 134 may be a connecting loop unimportant for overall structure.

We decided to also look for helices because of the report

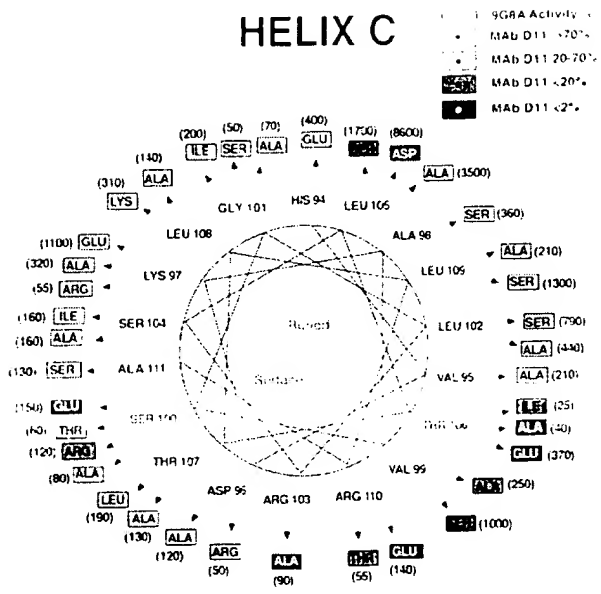


Fig 6. Helical wheel presentation of amino acids 94 to 111 (HELIX C). Amino acid residues (closest to the circle) represent those found in the human EPO sequence. Arrows point towards mutations (in boxes). Mutations that reduced D11 immunoreactivity are shaded as indicated. The effect of each mutation on 9G8A immunoreactivity is shown in parentheses. 9G8A immunoreactivity values are shown as percent of total antigen present as determined by immunoassay (see text). The predicted buried and exposed surfaces of the helix are shown.

that rHuEPO is 50% α-helix.[5] We reasoned that helical regions in rHuEPO would be amphipathic. That is, there would be buried and exposed surfaces on the helices. Therefore, mutations in residues on the buried surface would likely affect rHuEPO structure whereas similar mutations on the exposed surface would not. This would be especially true for mutations that introduced polar residues into buried positions. The effect of mutations on rHuEPO structure was monitored by 9G8A immunoreactivity as previously described.

The entire length of rHuEPO was examined by plotting the immunoassay data for rHuEPO variants on helical wheels.[22] The plots were then examined to identify regions that concentrated mutations that resulted in high 9G8A immunoreactivity on one side of the wheel. Three potential helices were identified using this method (residues 62 through 81, 96 through 110 and 144 through 157; Figs 5, 6, and 7).

A property of amphipathic helices is that the buried surface contains predominantly hydrophobic residues and the exposed surface contains predominantly hydrophillic residues. As shown in Figs 5, 6, and 7, each of the three regions containing the putative helices can be divided into surfaces containing predominantly polar or hydrophobic residues. Furthermore, the mutations in residues on the hydrophobic side of each wheel resulted in increases in 9G8A immunoreactivity. In contrast, polar to nonpolar mutations in the predicted solvent exposed region did not result in increases in 9G8A immunoreactivity. These results are consistent with α-helical structure for these three regions.

Support for the existence of Helix B (aas 62-81) comes from the D11 epitope mapping data (Figs 2 and 5). Residues directly contacting the D11 antibody should be on the exposed surface of rHuEPO and mutations in them should result in large decreases in immunoreactivity. In contrast, if residues in a helix are buried but near to residues that are part of the epitope then the immunoreactivity to variants containing these mutations should decrease but to a lesser extent because of nonspecific effects on secondary and tertiary structure. Examination of Helix B shows that the putative exposed surface contains the residues that appear to be most important for D11 immunoreactivity, Trp^{64}, Glu^{72} and Leu^{75}. Mutations in these residues resulted in 10-fold, 25-fold, and 30-fold decreases, respectively, in D11 immunoreactivity, but had normal 9G8A immunoreactivity. Thus, these three amino acids may be directly involved in D11 binding and are therefore likely to be on the exposed surface of the molecule. Mutations in residues on the putative buried surface also resulted in decreased D11 immunoreactivity but to a lesser degree. In addition, the D11 immunoreactivity was inversely related to 9G8A immunoreactivity suggesting that the decreased D11 immunoreactivity is a consequence of unfolding of rHuEPO. For example see residues 67, 70, and 74.

Additional analysis of helix B allows prediction of possible amino terminal and carboxy terminal helix caps. Resi-

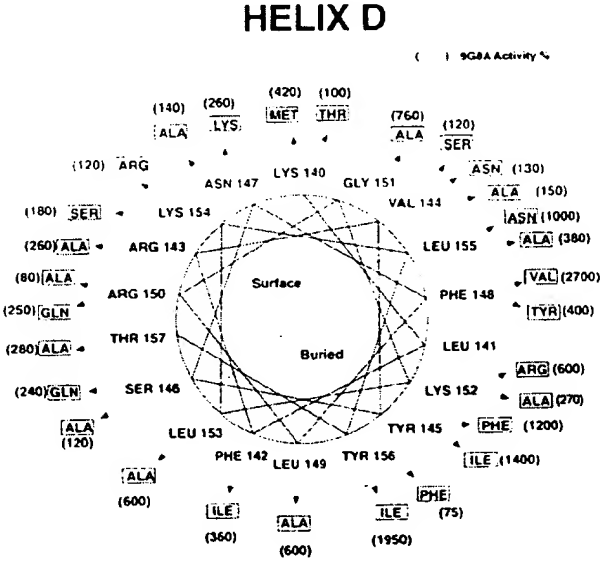


Fig 7. Helical wheel presentation of amino acids 140 to 157 (HELIX D). Amino acid residues (closest to the circle) represent those found in the human EPO sequence. Arrows point towards mutations (in boxes). The effect of each mutation on 9G8A immunoreactivity is shown in parentheses. 9G8A immunoreactivity values are shown as percent of total antigen present as determined by immunoassay (see text). The predicted buried and exposed surfaces of the helix are shown.

dues amino terminal to residue 58 and carboxy terminal to 76 lose amphipathic structure according to helical wheel plots of 9G8A immunoreactivity. For example, if the helix extended beyond Arg[76], then Ala[77] is predicted to be on the exposed surface. However, introduction of a Ser residue at position 79 results in increased 9G8A immunoreactivity (Table 1). In any case, it is unlikely that the helix extends beyond Arg[76] because of the Gly at 77. Glycine residues are thought to destabilize helices and are often found at helix ends.[25] In addition, positively charged residues are often found near the α-helix C-terminus and acidic residues are often found near the amino terminus because of their stabilizing effect on the helix dipole.[25] This would be consistent with helix caps at either Glu 58 or Glu[57] and a C-terminal cap at Arg76.

As shown in Fig 6 residues between 96 and 110 also have a helical wheel pattern consistent with an α-helix. As with helix B, mutations that result in increased 9G8A immunoreactivity are concentrated on one side of helix C. A Gly at position 112 suggests that the helix stops prior to this position. Thus, Arg[110] is a possible C-terminal helix cap. A possible amino terminal helix cap is Asp[96]. Helix caps at Asp[96] and Arg[110] would result in a helix the same length as helix B. In support of this proposed helix is the observation that Lys[97] is thought to be partially buried because it reacts, but poorly, with reagents that react with Lys amino groups.[26] The 9G8A immunoassay results (Fig 6) suggest that this residue may be at the buried/surface interface. The D11 epitope mapping results are also consistent with a helix in this region. As with helix B the mutations that had the greatest effect on immunoreactivity are concentrated on the putative exposed surface (Ser[100], Arg[103], Thr[106], and Arg[110]). Mutations in these amino acids result in over 100-fold decreases in D11 immunoreactivity. The exception is mutations at Leu[105]. However, variants with mutations at Leu[105] are secreted poorly and have low RIA-P immunoreactivity suggesting that the mutations at this position affect folding (eg, see Ser[105] Table 1). Other mutations in the putative buried surface also affect D11 immunoreactivity. However, in all cases these mutations are in hydrophobic amino acids and they result in increased 9G8A immunoreactivity. Therefore it is likely that the reduced immunoreactivity is because of nonspecific effects on rHuEPO folding.

Residues 144 to 157 also have a helical wheel pattern consistent with an α-helix. In a manner similar to helices B and C, mutations that increase 9G8A immunoreactivity are concentrated on the hydrophobic side of the helix. There is one charged residue on the predicted buried surface, Lys[152]. Mutations in Lys[152] resulted in increased 9G8A immunoreactivity however (Fig 7). In addition this residue is thought to be buried according to its poor reactivity with chemical modifying reagents.[6,24] The location of the helix caps is unclear. However it is unlikely that the helix extends beyond Gly[158], a helix breaking residue. Possible amino terminal ends are Lys[140] and Val[141]. A helix cap starting at Val[141] would result in a helix the same length as those predicted for helices B and C. It is unlikely that the helix begins before residue 138 because the F12 epitope mapping data suggests that 138 is at or near to the F12 epitope. If the helix began before 138 it would not be possible to position this residue near to the other residues thought to be part of the epitope, residues 32 and 88.

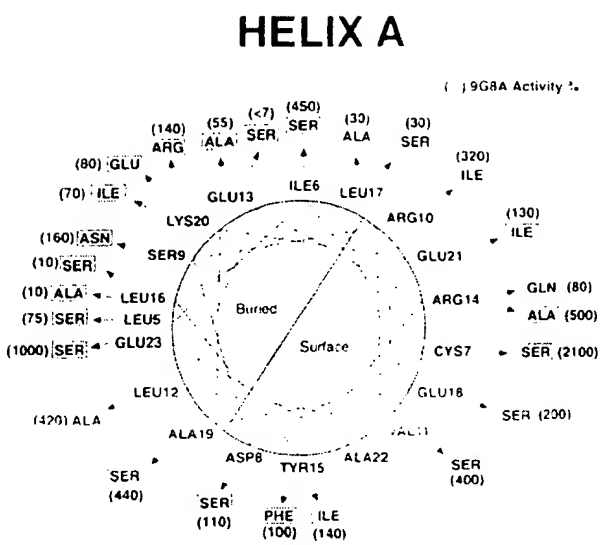


Fig 8. Helical wheel presentation of amino acids 5 to 23 (HELIX A). Amino acid residues (closest to the circle) represent those found in the human EPO sequence. Arrows point towards mutations (in boxes). The effect of each mutation on 9G8A immunoreactivity is shown in parentheses. 9G8A immunoreactivity values are shown as percent of total antigen present as determined by immunoassay (see text). The predicted buried and exposed surfaces of the helix are shown.

Human EPO has been reported to fall into the 4-helix bundle category of cytokine folds.[6,7,9] Therefore there should be a fourth, amino terminal, helix. We plotted this region on a helical wheel (Fig 8). Hydrophobic residues are generally concentrated on one side of the potential helix: Leu[7], Ile[8], Leu[12], Leu[15], and Leu[17]. However, a helical wheel plot does not concentrate mutations that result in increased 9G8A immunoreactivity on any one side. A complication with this analysis is that the 9G8A epitope maps to this region. In fact, rHuEPO with mutations in some of the putative buried residues, Glu[13], Leu[15], and Leu[17], all have reduced apparent 9G8A immunoreactivity (previously described). If these residues are indeed buried, it would explain the enhanced immunoreactivity to rHuEPO with 9G8A upon unfolding. There is support for a helix in this region as well as possible orientation of its exposed and buried surfaces. Tyr[15] is readily iodinated suggesting its surface location.[27] In addition, Arg[14] is thought to be part of the active site.[28] This would put this residue on the surface as well. Finally, two mutations on the predicted buried surface, Leu[12] to Ala and Ala[19] to Ser mutations resulted in high 9G8A immunoreactivity suggesting that they are buried. Possible helix caps are Arg[4], or Asp[5], and Lys[20].

A model consistent with the antibody epitope mapping is a 4-helix bundle. We have developed a model of rHuEPO structure using the data reported here. The following constraints were used. Residues 7 and 161 were joined by a

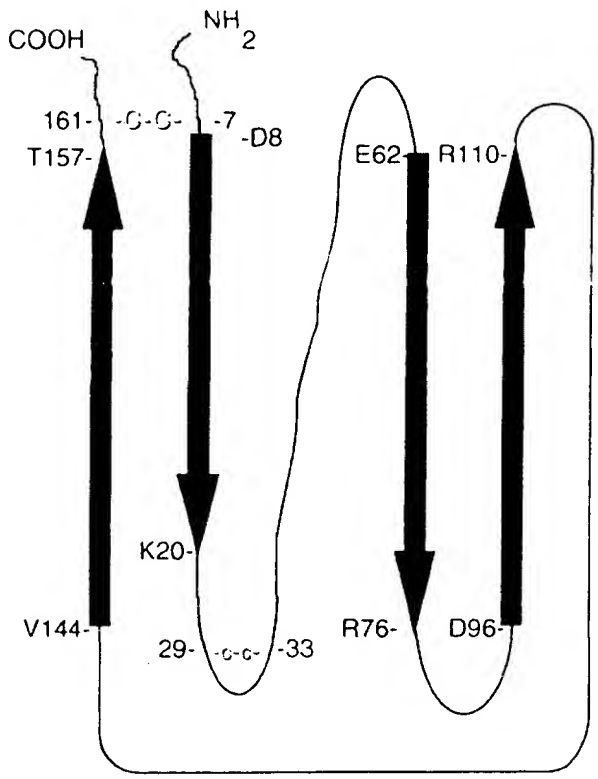


Fig 9. Modeling of EPO structure. The model is based on information described in the text. Arrows oriented in the direction shown represent potential α-helices labeled A, B, C, and D. Lines represent connecting loops. Numbers represent locations of the indicated amino acid residues where 1 is the amino terminus of rHuEPO following removal of the signal peptide. Disulfide bonds are indicated (C-C).

disulfide bond. Residues 29 to 33 were assumed to contain a turn because of the Cys^{29}-Cys^{33} bond. F12 epitope mapping results suggested that residues 29 through 33, 86 through 88, and 138 were positioned close to each other. D11 epitope mapping results suggested that residues 64, 72, 75, 100, 103, and 106 were also close to each other. 9G8A mapping results indicated that its epitope includes residues 13, 16, and 17. The results from the accompanying paper[11] suggested that MoAbs F12, D11, and 9G8A recognize nonoverlapping epitopes. Therefore residues in each of the three epitopes were separated. This data allowed prediction of an overall protein fold. We then superimposed on the model the predicted helices (8 through 20, 62 through 76, 96 through 110, and 144 through 157) and connecting loop (114 through 134).

The model of rHuEPO structure is shown in Fig 9. This structure includes 4 closely packed helices labeled A, B, C, and D with a up-up-down-down helix topology. Long loops, 33 through 62 and 110 through 140, connect helices A to B and C to D, respectively. The residues important for protein stability, 42, 70, 105, 148, 152 and 156, are positioned in areas consistent with their being part of a hydrophobic core. The D11 epitope included helices B and C. The 9G8A epitope is in the A helix opposite from the D11 epitope. The F12 epitope includes the potential turns on the bottom of the structure.

DISCUSSION

We describe here an immunological approach to determination of rHuEPO structure. Crucial to this approach is the availability of individual rHuEPO point mutations and the ability to determine the effects of the mutations on folding and immunoreactivity. Five different immunoassays were used for this effort, including 3 different monoclonals and 2 different polyclonals. In addition 2 different secondary antibodies were used in ELISAs. The 3 monoclonals recognized nonoverlapping epitopes which reduced the possibility that mutations would reduce immunoreactivity for all 3 simultaneously. The combination of all these assays resulted in a high probability that at least one could be found to estimate rHuEPO concentration.

The predicted structure of rHuEPO based on the epitope mapping and 9G8A folding experiments is a 4-helix bundle (Fig 9). There is other data consistent with this model. First, mouse and human EPO sequences are highly conserved[1,23,25] and both molecules stimulate human and mouse erythropoiesis.[26] This suggests that the structures of the two molecules are similar, particularly in the receptor binding regions. Mouse EPO has Cys^{33} to Pro and Arg^{139} to Cys substitutions as compared to human EPO. It has been proposed that in mouse EPO Cys^{29} and Cys^{139} form a disulfide bond. The model described here positions residues 29 and 139 close together. Second, amino acid residues 32 to 36, 53 to 57, 78 to 82, and 111 to 119 can be deleted yet the protein can maintain its structure.[7] We have also found that residues in the 111 through 139 region can be mutated without destabilizing the structure. Loop regions can often tolerate insertions and deletions while helical and beta sheet regions do not.[12-14] In agreement with our model, these amino acids are predicted to be part of either turns or interconnecting loops. Third, crosslinking studies indicate that certain residues are close together.[6] These include the amino terminus and Lys^{116}, Lys^{45} and Lys^{140}, Lys^{52} and Lys^{140}, and Lys^{52} and Lys^{154}. In addition the poor reactivity of Lys^{152} suggests that it is buried. These crosslinked residues are all positioned relatively close to each other on our model. Finally, active site residues, Arg^{14}, Arg^{103}, Ser^{104}, and Gly^{101}[26,30] are positioned on predicted surfaces and generally concentrate in patches on the model.

Several different models of EPO structure have been proposed elsewhere, all based on 4-helix bundles.[6-9] The models proposed by Bazan,[9] Boissel et al,[7] and Hanui et al[8] are all similar to the model proposed here. However they all differ in the location of helices and turns. The Bazan model was developed according to structural homology to human growth hormone. It predicts longer helices than those proposed here. The Boissel model was developed using second-

ary structure predictive methods and computer modeling. This model was consistent with deletion mutant data.[7] Both models are consistent with some but not all of our data. For example, the Boissel model positions residue 138 distant from residues 32 and 88. Epitope mapping with MoAb F12 positions these regions close together.

It has been reported that EPO neutralizing antibodies inhibit binding of EPO to its receptor.[31] The D11 antibody described here is neutralizing. This suggests that its epitope, residues 99 through 110 and 64 through 79, may include the receptor binding domain. Sytowski and Donahue[32] also identified neutralizing antibodies that were raised against synthetic peptides. This study showed that antibodies raised against peptides 99 through 118 and 111 through 129 would neutralize EPO activity. In vitro mutagenesis studies indicate further that the 100 through 110 region is important for biological activity but the 64 through 79 region is not.[26,30] This is consistent with our observations. The 9G8A antibody epitope maps to residues 13, 16, and 17. It did not neutralize rHuEPO activity.[11] Sytowski et al also raised antipeptide antibodies to residues 1 through 26 that were reported to be nonneutralizing.[32] In contrast Egrie et al[15] raised a neutralizing antibody to an amino terminal (residues 1 through 20) peptide. In addition, Arg^{14} has been reported to be part of the active site.[26] These results suggest a role for the amino terminal region in biological activity. Therefore, the N-terminal region active site residues involved in receptor binding must be positioned in such a way that they would not be stearically hindered by 9G8A. Indeed, the 9G8A epitope may be on the buried surface of the helix. Another possibility is that the 9G8A antibody is unable to neutralize rHuEPO activity because the affinity of 9G8A to its epitope is too low to block receptor binding when rHuEPO is properly folded. Additional information on the location of the rHuEPO active site may resolve the issue.

ACKNOWLEDGMENT

We thank B. Aguero, J. Grant, D. Greene, and X. Wang for expert technical assistance; J. Egrie for technical assistance and helpful discussions, K. Chen, S. Suggs, A. Harcourte, J. Katzowitz, and Ann Janssen for DNA sequence analysis and T. Jones, E. Fisher and Burt Goodman for synthetic peptides and oligonucleotides.

REFERENCES

1. Lin F-K, Suggs S, Lin C-H, Browne JK, Smalling R, Egrie JC, Chen KK, Fox GM, Martin F, Stabinski Z, Badrawi SM, Lai P-H, and Goldwasser E: Cloning and expression of the human erythropoietin gene. Proc Natl Acad Sci USA 82:7580, 1985

2. Jacobs K, Shoemaker C, Rudersdorf R, Neill SD, Kaufman RJ, Mufson A, Seehra J, Jones SS, Hewick R, Fritsch EF, Kawakita M, Shimizu T, Miyake T: Isolation and characterization of genomic and cDNA clones of human erythropoietin. Nature 313:806, 1985

3. Browne JK, Cohen AM, Egrie JC, Lai P-H, Lin F-K, Strickland T, Watson E, Stebbing N: Erythropoietin: Gene cloning, protein structure, and biological properties. Cold Spring Harbor Symp Quant Biol 51:693, 1986

4. Egrie JC, Strickland TS, Lane J, Aoki K, Cohen AM, Smalling R, Trail G, Lin F-K, Browne JK, Hines DK: Characterization and biological effects of recombinant human erythropoietin. Immunobiology. 172:213, 1986

5. Davis JM, Arakawa T, Strickland TW, Yphantis DA: Characterization of recombinant human erythropoietin produced in Chinese hamster ovary cells. Biochemistry 26:2633, 1987

6. Haniu M, Narhi LO, Arakawa T, Elliott S, Rohde MF: Recombinant human erythropoietin (rHuEPO): Cross-linking with disuccinimidyl esters and identification of the interfacing domains in EPO. Protein Sci 2:1441, 1993

7. Boissel J-P, Lee WR, Presnell SR, Cohen FE, Bunn HF: Erythropoietin structure–function relationships, mutant proteins that test a model of tertiary structure. J Biol Chem 268:15983, 1993

8. Goldwasser E: Structure-function relationships of erythropoietin, in Erslev AJ, Adamson JW, Eschbach JW, Wincarls CG (eds): Erythropoietin, Molecular, Cellular and Clinical Biology. Baltimore, MD, Johns Hopkins University Press, 1991, p 41

9. Bazan JF: Haemopoietic receptors and helical cytokines. Immunol Today 11:350, 1990

10. Jin L, Cohen FE, Wells JA: Structure from function: Screening structural models with functional data. Proc Natl Acad Sci USA 91:113, 1994

11. Elliott S, Boone T, Chang D, Delorme E, Dunn C, Egrie J, Giffin J, Lorenzini A, Talbot C, Hesterberg L: Isolation and characterization of confonnation sensitive anti-erythropoietin monoclonal antibodies: Effect of disulfide bonds on rHuEPO structure. Blood 87:2714, 1996

12. Freimuth PI, Taylor JW, Kaiser ET: Introduction of guest peptides into *Escherichia coli* alkaline phosphatase. J Biol Chem 265:896, 1990

13. Pakula AA, Sauer RT: Genetic analysis of protein stability and function. Annu Rev Genet 23:289, 1989

14. Bowie JU, Reidhaar-Olson JF, Lim WA, Sauer RT: Deciphering the message in protein sequences: Tolerance to amino acid substitutions. Science 247:1306, 1990

15. Egrie J, Lane J: Use of monoclonal and polyclonal antipeptide antibodies to assay and characterize erythropoietin. Fed Proc 43:1892, 1984

16. Law ML, Cai G-Y, Lin F-K, Wei Q, Huang S-Z, Hartz JH, Morse H, Lin C-W, Jones C, Kao F-T: Chromosomal assignment of the human erythropoietin gene and its DNA polymorphism. Proc Natl Acad Sci USA 83:6920, 1986

17. Yanisch-Perron C, Viera J, Messing J: Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103, 1985

18. Kunkel TA, Roberts JD, Zakour RA: Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol 154:367, 1987

19. Messing J: New m13 vectors for cloning. Methods Enzymol 101:20, 1983

20. Wigler M, Pellicer A, Silverstein S, Axel R: Biochemical transfer of single-copy eucaryotic genes using total cellular DNA as donor. Cell 14:725, 1978

21. Lin F-K, Lin C-H, Lai P-L, Browne JK, Egrie JC, Smalling R, Fox GM, Chen KK, Castro M, Suggs S: Monkey erythropoietin gene: Cloning, expression and comparison with the human erythropoietin gene. Gene 44:201, 1986

22. Schiffer M, Edmundson AB: Use of helical wheels to represent the structures of proteins and to identify segments with helical potential. Biophys J 7:121, 1967

23. Richardson JS, Richardson DC: Principles and patterns of protein conformation, in Fasman GD (ed): Prediction of Protein Structure and the Principles of Protein Conformation. New York, NY, Plenum, 1989, p 1

24. Satake R, Kozutsumi H, Takeuchi M, Asano K: Chemical modification of erythropoietin: An increase in in vitro activity by guanidination. Biochim Biophys Acta 1038:125, 1990

25. Feldman L, Chung T, Zurbuch D, Donahue KA, Powell JS, Sytkowski AJ: Identification of a structural role for tyrosine-145 in Human erythropoietin. Blood 74:194a, 1989 (abstr)

26. Wen D, Biossel J-P, Showers M, Ruch BC, Bunn HF: Erythropoietin structure-function relationships; Identification of functionally important domains. J Biol Chem 269:22839, 1994

27. McDonald JD, Lin F-K, Goldwasser E: Cloning, sequencing and evolutionary analysis of the mouse erythropoietin gene. Mol Cell Biol 6:842, 1986

28. Shoemaker CB, Mitsock LD: Murine erythropoietin gene: Cloning expression, and human gene homology. Mol Cell Biol 6:849, 1986

29. Cotes PM, Bangham DR: Bio-assay of erythropoietin in mice made polycythaemic by exposure to air at reduced pressure. Nature 191:1065, 1961

30. Grodberg J, Davis KL, Sytowski AJ: Alanine scanning mutagenesis of human erythropoietin identifies four amino acids which are critical for biological activity. Eur J Biochem 218:597, 1993

31. D'Andrea AD, Szklut PJ, Lodish HF, Alderman EM: Inhibition of receptor binding and neutralization of bioactivity by anti-erythropoietin monoclonal antibodies. Blood 75:874, 1990

32. Sytowski AJ, Donahue KA: Immunochemical studies of human erythropoietin using site-specific anti-peptide antibodies. J Biol Chem 262:1161, 1987

Express Mail Label No. EV399139625US

Erythropoietin Structure-Function Relationships: High Degree of Sequence Homology Among Mammals

By Danyi Wen, Jean-Paul R. Boissel, Timothy E. Tracy, Robert H. Gruninger, Linda S. Mulcahy, John Czelusniak, Morris Goodman, and H. Franklin Bunn

To investigate structure-function relationships of erythropoietin (Epo), we have obtained cDNA sequences that encode the mature Epo protein of a variety of mammals. A first set of primers, corresponding to conserved nucleotide sequences between mouse and human DNAs, allowed us to amplify by polymerase chain reaction (PCR) intron 1/exon 2 fragments from genomic DNA of the hamster, cat, lion, dog, horse, sheep, dolphin, and pig. Sequencing of these fragments permitted the design of a second generation of species-specific primers. RNA was prepared from anemic kidneys and reverse-transcribed. Using our battery of species-specific 5' primers, we were able to successfully PCR-amplify Epo cDNA from Rhesus monkey, rat, sheep, dog, cat, and pig. Deduced amino acid sequences of mature Epo proteins from these animals, in combination with known sequences for human, Cynomolgus monkey, and mouse, showed a high degree of homology, which explains the biologic and immunological cross-reactivity that has been observed in a number of species. Human Epo is 91% identical to monkey Epo, 85% to cat and dog Epo, and 80% to 82% to pig, sheep, mouse, and rat Epos. There was full conservation of (1) the disulfide bridge linking the NH_2 and COOH termini; (2) N-glycosylation sites; and (3) predicted amphipathic α-helices. In contrast, the short disulfide bridge (C29/C33 in humans) is not invariant. Cys33 was replaced by a Pro in rodents. Most of the amino acid replacements were conservative. The C-terminal part of the loop between the C and D helices showed the most variation, with several amino acid substitutions, deletions, and/or insertions. Calculations of maximum parsimony for intron 1/exon 2 sequences as well as coding sequences enabled the construction of cladograms that are in good agreement with known phylogenetic relationships.

ERYTHROPOIETIN (Epo) is the hematopoietic cytokine that regulates red blood cell production. In mammals, this 34-Kd glycoprotein hormone is produced in the fetal liver and adult kidney, circulates in the bloodstream, and binds to receptors on committed progenitor cells in the bone marrow and other hematopoietic tissues, resulting in proliferation and terminal maturation of erythroid cells.[1] The expression of both Epo mRNA and Epo protein is markedly increased by hypoxia, owing to a 3' enhancer and to highly conserved elements in the promoter region.[2-5] This elegant servomechanism enables Epo to regulate the red blood cell mass of humans and other animals.

A large number of early physiologic studies have established extensive, although incomplete, biologic cross-reactivity between the Epos of humans and a number of other mammals, including the mouse, rat, sheep, and rabbit.[1,6-8] In contrast, nonmammalian vertebrates (amphibians,[9,10] birds,[11,12] and fish[13]) have erythropoietic hormones that fail to cross-react with mammalian erythroid cells, and vice-versa.[9-11]

Thus far, the Epo genes of a human,[14,15] a monkey (Macaca fascicularis),[16] and a rodent (the mouse[17,18]) have been cloned, sequenced, and expressed. In view of the marked cross-reactivity between mammalian Epos, it is not surprising that there is a high degree of sequence homology in the coding region of the mature secreted proteins. In keeping with their close phylogenetic relationship, human and monkey Epos are 94% and 91% identical in nucleotide and amino acid sequence, respectively. In contrast, human and mouse Epos are 76% identical in nucleotide sequence and 80% identical in amino acid sequence.

Direct information on the three-dimensional structure of Epo is not yet available. Insights into structure-function relationships of Epo can be gained from the analyses of a more complete set of animal sequences. Such information could be useful for sequence-based computer modeling of three-dimensional structure. Moreover, a larger data base would permit the identification of highly conserved domains that are likely to be crucial to folding and/or biologic function. Finally, comparative amino acid and nucleotide sequence information on Epo provides additional data for investigating phylogenetic relationships.

We report here the use of polymerase chain reaction (PCR)-based techniques for obtaining the full coding sequences of Epos of the Rhesus monkey, rat, sheep, pig, dog, and cat, as well as partial sequences from three other species. The design and preparation of mutants that test a 4 α helical bundle model of Epo tertiary structure have been presented elsewhere.[19]

MATERIALS AND METHODS

Animal Samples

Human hepatoma Hep3B cell line and multiple kidney-derived cell lines from hamster (BHK), sheep (MDOK), pig (LLC-PK1), dog (MDCK), cat (CRF-K), lion (PAL1-K), and spotted dolphin (SP1-K) were obtained through the American Type Culture Collection (Rockville, MD). Monolayer cells were grown in 100 × 20 mm tissue culture dishes using recommended media and maintained in

From the Hematology/Oncology Division, Brigham and Women's Hospital, Harvard Medical School, Boston, MA; the R.W. Johnson Pharmaceutical Research Institute, Raritan, NJ; and the Departments of Anatomy and Cell Biology and of Molecular Biology and Genetics, Wayne State University School of Medicine, Detroit, MI.

Submitted March 11, 1993; accepted May 12, 1993.

Supported by National Institutes of Health Grant No. R01 HL42949 (to H.F.B.), by NSF Grant No. DEB9116098 (to M.G.), and by the R.W. Johnson Pharmaceutical Research Institute.

Address reprint requests to H. Franklin Bunn, MD, Hematology/Oncology Research: LMRC 2, Brigham and Women's Hospital, Rm 223, 221 Longwood Ave, Boston, MA 02115.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. section 1734 solely to indicate this fact.

0006-4971/93/8205-0016$3.00/0

		Localization (bp)	
		human(1)	mouse(2)
IV1	5' TGAAGTTTGGCCGAGAAGTGGATGC 3'	984-1009	994-1018
EX2R	5' AAGA(T/G)GTACCTCTCCAG(A/G)ACTCG 3'	1296-1318	1302-1324
EX5	5' CTGCTCCACTCCGAACA(C/A)TCAC 3'	2651-2672	3183-3204
NCOI	5' CTGGAGTGTCCATGGGACAG 3'	2885-2904	3415-3434
ATG	5' AGGCGCGGAGATGGGGGTGC 3'	615-634	633-652

Fig 1. Oligonucleotide primer sequences. The localization of the primer coincides to previously published nucleotidic sequences for the human[15] (1) and murine[17] (2) Epo genes. EX2R and EX5 are not completely conserved between the human and mouse (respectively, 93% and 95% identity). An equal amount of each possible nucleotide was incorporated during the corresponding cycles of those primer syntheses. EX2R and NCOI are reverse primers and their sequences represent the antisense DNA strand.

a humidified 5% CO_2/95% air incubator at 37°C. In some experiments, cells were made hypoxic by overnight incubation in 1% O_2, 5% CO_2, and 94% N_2 at 37°C.

Poly A$^+$ RNA prepared from Rhesus monkey kidney was purchased from Clontech (Palo Alto, CA). Kidneys from anemic cat and horse were obtained from veterinarians at Tufts School of Veterinary Medicine (Grafton, MA) and the University of Nebraska (Lincoln, NE), respectively. Anemia was induced by three consecutive daily intraperitoneal injections of phenylhydrazine (60 mg/kg body weight) into male Sprague-Dawley rats (SD-strain) and by repeated bleeding of the dog, the sheep, and the pig. Kidneys were aseptically removed after induction and stored immediately in liquid nitrogen.

A 4-kb genomic human Epo clone (gEpo4) was provided by Genetics Institute (Boston, MA).

DNA and RNA Preparations

DNAs from cell lines or homogenized kidneys were extracted using pancreatic RNAase/SDS/proteinase K following a procedure modified from Blin and Stafford.[20]

Kidneys were homogenized in 4 mol/L guanidine isothiocyanate (10 mL/g of frozen organ) containing 0.1 mol/L β-mercaptoethanol and 2% N-lauryl-sarcosine. Total RNAs were isolated by centrifugation over 5.7 mol/L CsCl.[21] After ethanol precipitation, the samples were resuspended in diethyl pyrocarbonate-treated water and

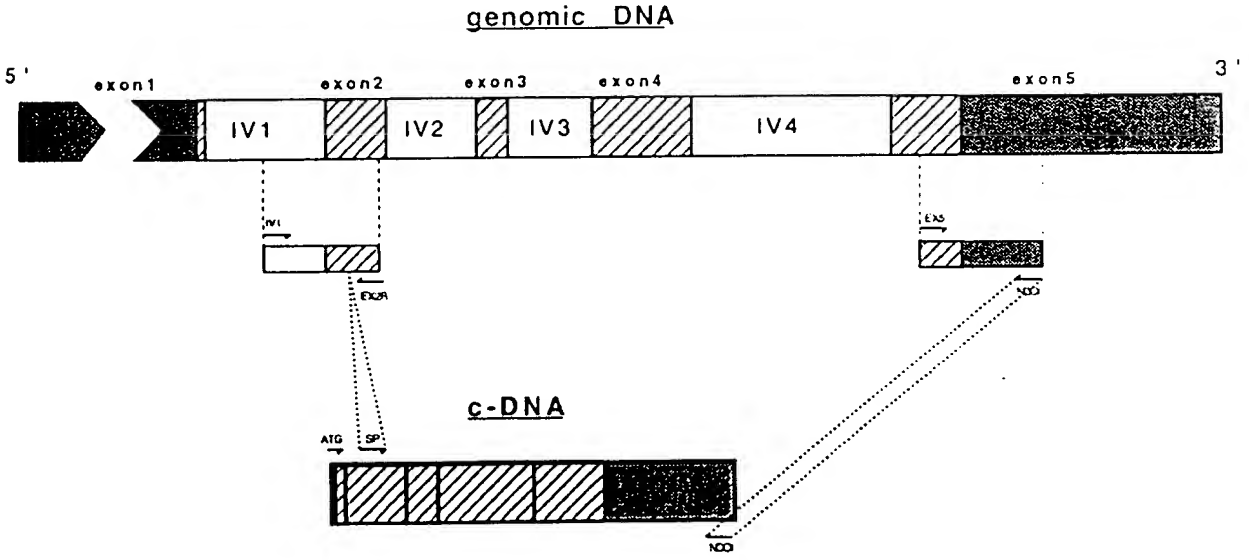


Fig 2. PCR strategy used for the cloning of mammalian cDNAs containing the complete coding sequence of the mature Epo protein. Genomic amplification and sequencing of exonic fragments, localized upstream and downstream from the nucleotide portion coding for the mature protein, allowed the design of species specific primers (SP). Use of those SP primers and/or of primers, 100% conserved between the human and mouse (5 ATG and 3 NCOI) on cDNA templates prepared from kidney of uninduced or hypoxia-induced animals, allows the amplification of a variety of mammalian Epo clones. Sense (→) and antisense (←) primers are represented by the arrows. Dashed boxes correspond to the coding part (propeptide and mature protein) of the five Epo exons. Shaded boxes represent the 5 and 3 untranslated regions.

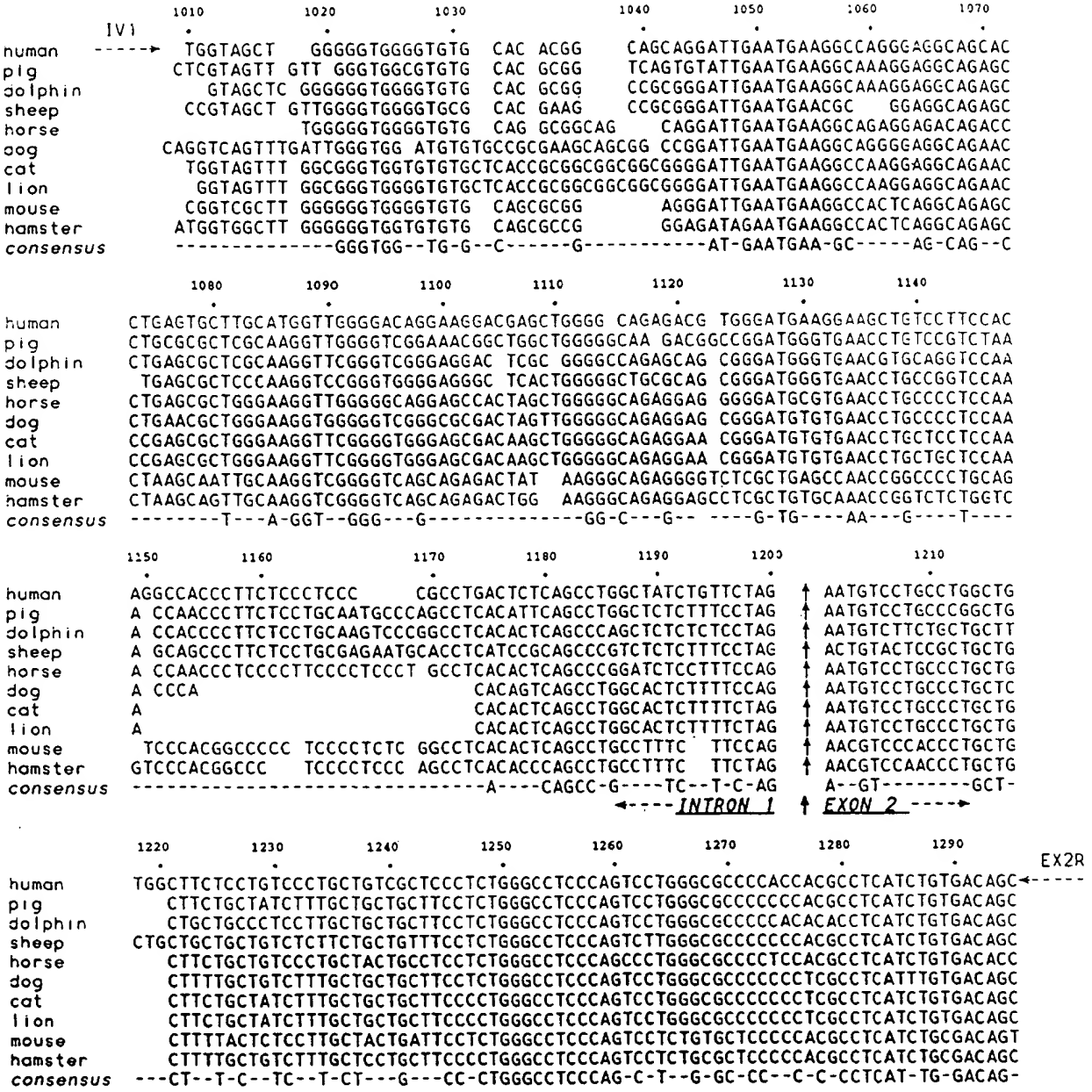


Fig 3. Comparison of IV1/EX2R sequences from various mammals. The numbering corresponds to the published human genomic sequence.[16] The reported human sequence was obtained from the amplification of purified genomic DNA from Hep3B cells and agreed with the sequence previously reported. The mouse sequence is from McDonald et al.[17] The horse sequence was obtained from kidney-extracted genomic DNA. All the other sequences were established from PCR amplification of genomic DNAs purified from several mammalian renal-derived cell lines. The consensus sequence indicates the positions in which a unique nucleotide was found in all of the reported species. The boundary between intron 1 and exon 2 is represented by the ascending arrows. The locations of the two PCR primers, IV1 and EX2R, are shown by the dashed arrows.

stored at −70°C. Confluent monolayer cells were washed twice with sterile phosphate-buffered saline and directly lysed in the 4 mol/L guanidine isothiocyanate solution. Total RNAs were isolated as described above for the kidneys.

RNA samples were first converted into single-strand cDNA. Two to 4 μg of total RNA from kidney or 500 ng to 1 μg of total RNA from cultured cells was denatured at 68°C in presence of 2 μg of oligo $dT_{(15)}$. Reverse transcription was performed in a 20 μL final volume, containing 50 mmol/L Tris-HCl, pH 8.3, 6 mmol/L $MgCl_2$, 40 mmol/L KCl, 10 mmol/L dithiothreitol (DTT), 500 μmol/L dNTPs, 20 U RNAsin (Promega, Madison, WI), and 20 U of avian myeloblastosis virus reverse transcriptase (AMV-SuperRT; Molecular Genetic Resources, Tampa, FL). The reaction was allowed to proceed at 37°C for 20 minutes and then at 42°C for 1 hour. Inactivation of the enzyme was performed at 100°C for 10 minutes. cDNA samples were stored at −70°C.

PCR

Twenty to 100 ng of genomic DNA or 0.5 to 1 μL of the RT-reaction were used as templates for amplification.

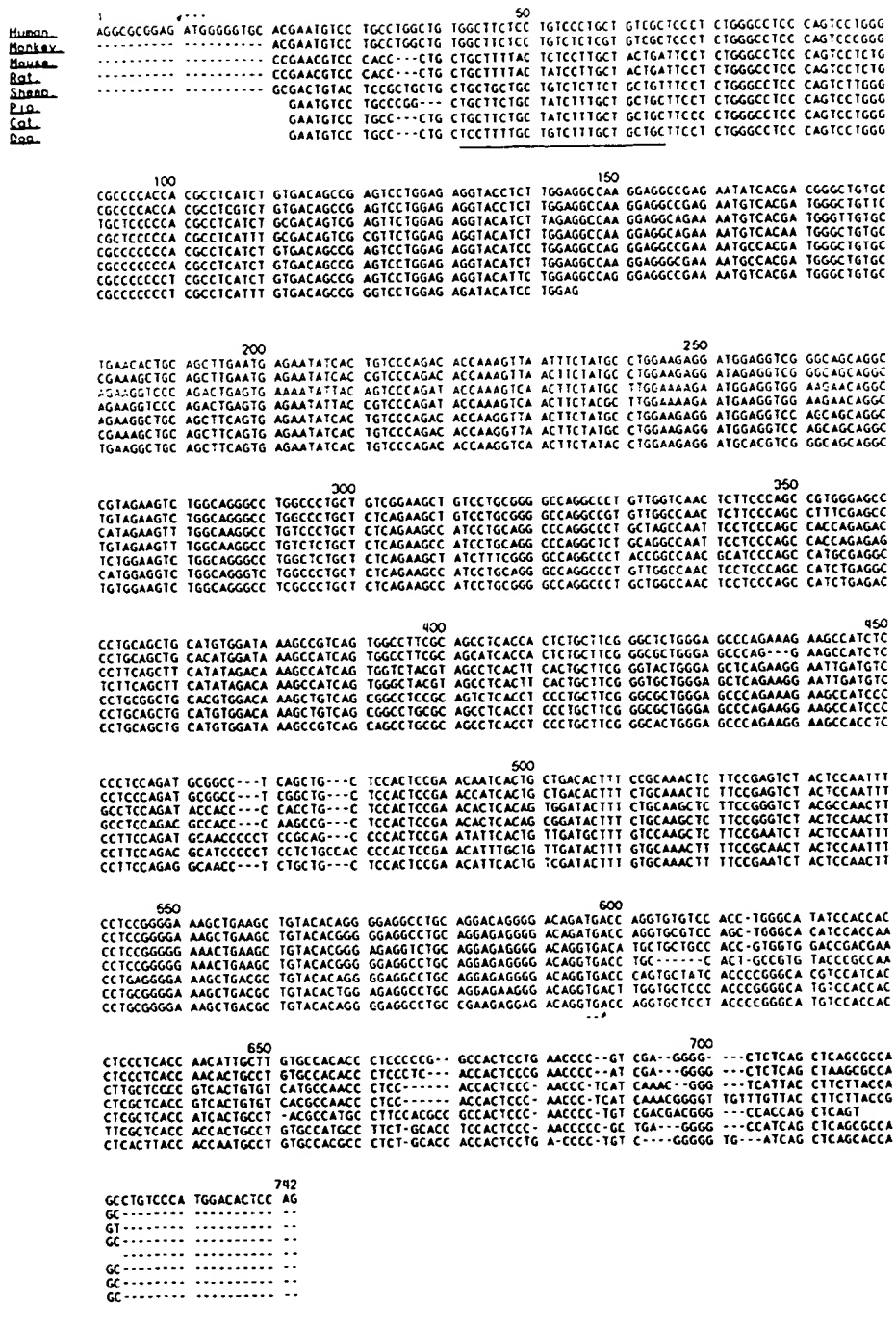


Fig 4. Alignment of the nucleotide sequences of mammalian Epo cDNAs. The mouse sequence corresponds to the one previously reported.[17] PCR-produced human sequence was obtained from hypoxia-induced Hep3B cell line and is in total agreement with the sequence previously published.[14,16] Monkey (*Macaca mulatta*) (GenBank; accession no. L10609), rat (L10608), sheep (L10610), pig (L10607), cat (L10606), and dog (L13027) were amplified from kidney-purified cDNAs. Amplifications of the human, monkey, rat, and sheep were realized using the ATG (residues 1 to 20) and NCOI (residues 723 to 742) primers. The line under residues 41 to 56 represents the specific SP1 primer used for the cloning of the pig, cat, and dog sequences. For these three mammals, sequences from SP1 are derived from the data presented in Fig 3. Arrows indicate the start and the end of the coding sequence (propeptide and mature protein).

Standard PCR reactions were performed in a 100 μL volume containing 50 mmol/L Tris-HCl, pH 8.3, 50 mmol/L KCl, 1.5 mmol/L $MgCl_2$, 0.01% wt/vol gelatin, 200 μmol/L dNTPs, 30 pmol/L of each sense (5') and antisense (3') primers, and 0.5 U of Taq DNA polymerase (Perkin Elmer Cetus, Emeryville, CA). After an initial denaturation step for 5 minutes at 95°C, routinely 35 cycles of PCR were performed on the DNA Thermal Cycler (Perkin Elmer Cetus). A denaturation step was performed at 94°C for 1 minute. For each sample and/or the primer pair used, annealing temperatures were optimized (from 45°C to 60°C) for a 2-minute reaction. A 3-minute elongation step at 72°C ended each cycle.

Amplification products were analyzed on 1% to 3% agarose gel-Tris-acetate-EDTA (TAE). Nu Sieve or Seaplaque agaroses (FMC, Rockland, ME) were used for preparative purposes. Specific PCR products were recovered from the gel by the use of the Gene Clean II kit (Bio 101, La Jolla, CA), by Spin X centrifugation (Costar, Cambridge, MA), or by standard phenol/chloroform extractions of melted gel slices.

Subcloning and Sequencing

Small (250 to 330 bp) amplified genomic fragments were directly subjected to automated sequencing, using the PCR primers as DNA sequencing primers.

	-27																										1		
Human	M	G	V	H	E	C	P	A	W	L	W	L	L	L	S	L	L	S	L	P	L	G	L	P	V	L	G↑	A	P	P
Monkey	M	G	V	H	E	C	P	A	W	L	W	L	L	L	S	L	V	S	L	P	L	G	L	P↑	V	P	G	A	P	P
Rodents	M	G	V	P	E	R	P	T	-	L	L	L	L	L	S	L	L	L	I	P	L	G	L	P	V	L	C	A	P	P
Sheep	M	G	V	R	D	C	T	P	L	L	L	L	L	L	S	L	L	L	F	P	L	G	L	P	V	L	G	A	P	P
Pig					E	C	P	A	R	-	L	L	L	L	S	L	L	L	L	P	L	G	L	P	V	L	G	A	P	P
Horse					E	C	P	A	-	L	L	L	L	L	S	L	L	L	P	P	L	G	L	P	A	L	G	A	P	P
Carnivores					E	C	P	A	-	L	L	L	L	L	S	L	L	L	L	P	L	G	L	P	V	L	G	A	P	P
Dolphin					E	C	-	L	L	L	L	L	L	P	S	L	L	L	L	P	L	G	L	P	V	L	G	A	P	P

Fig 6. Predicted amino acid sequences of Epo propeptide. The amino acid sequences are derived from data obtained from both genomic IV1/EX2R and cDNA amplifications. The numbering corresponds to the human sequence. Ala 1 is the first amino acid of the human mature protein. The ascending arrows show the site of the cleavage by the signal peptidase, as determined for the human and cynomolgus monkey proteins.

PCR-generated cDNA products were blunt-ended using Klenow fragment and 5′ phosphorylated by T4 DNA kinase (Boehringer Mannheim, Indianapolis, IN). They were then cloned into the *Sma* I site of the phagemid pBluescript II SK (Stratagene, La Jolla, CA). The double-stranded plasmid was sequenced using the SK and KS sequencing primers.

In all cases, sequencing reactions were performed on an Applied Biosystems 373A Automated DNA Sequencer using the DyeDeoxy Terminator Sequencer kit (Applied Biosystems, Foster City, CA) and thermal cycling with Taq DNA polymerase (Promega), as previously described.[22] To avoid errors, for each species, samples derived from 4 to 8 different amplifications were isolated and subcloned. In each case, both strands were sequenced. A very low error rate was encountered, in keeping with the 100% agreement between the rat cDNA sequence (see Fig 4) and that recently reported by Nagao et al.[23]

Mammalian Expression

Inserts containing full-length mammalian Epo coding sequences were subcloned into pSG5 plasmid (Stratagene) and were transiently expressed in Cos7 cells. Seventy-two-hour supernatants were tested for their ability to sustain cellular proliferation of the Epo-dependent HCD57 cell line.[24]

RESULTS AND DISCUSSION

Because of the known biologic cross-reactivity between various mammalian Epos and the presumption of strong homology in the coding sequences, a logical cloning strategy would be to screen cDNA libraries with moderate to low stringency. However, Epo mRNA is expressed at barely detectable levels in all tissues except in hypoxic kidney and liver.[25-28] Therefore, it would be necessary to generate libraries for each of the species of interest. Our primary research interest is in structure-function relationships of Epo and, therefore, we require full coding sequences of the mature Epo protein, rather than full cDNA sequences. Accordingly, we elected to use a PCR cloning strategy predicated on known strong homologies in Epo cDNA sequences that flank the region encoding the mature protein. There was sufficient conservation of sequence around the NCO1 site in the 3′ untranslated region (UTR) that a single primer could be used to amplify full-length mature Epo 3′ coding sequences in the rat, sheep, pig, and cat as well as in the dog. However, we were unsuccessful in finding a universal 5′ primer and therefore had to generate species-specific sequences from amplifications of genomic DNA. This two-stage PCR strategy proved to be satisfactory in generating highly accurate sequence information with a minimum expenditure of time and resources. In all instances, there was full agreement between sequences of complementary strands.

Preparation of Mammalian Epo cDNAs

Primer design for PCR amplification of Epo. The sequences of Epo genes from three species have been already published: two primates, the human[14,15] and Cynomolgus monkey (*Macaca fascicularis*),[16] and a rodent, the mouse (*Mus musculus*).[17,18] There is substantial homology at the nucleotide level not only in the coding sequences, but also in portions of introns and in 5′ and 3′ untranslated sequences of exons 1 and 5, respectively. A number of primers, corresponding to conserved nucleotide sequences between the mouse and human, were synthesized and tested for their ability to produce genomic PCR fragments from a wide variety of mammalian cell lines. A 4-kb human genomic clone and genomic DNA extracted from the human hepatoma cell line Hep3B were used as control templates for the PCR amplifications.

In particular, two pairs of primers, IV1/EX2R and EX5/NCOI (Fig 1), directed the amplifications of fragments of 330 and 250 bp, respectively, from genomic DNA of all human and mammalian cell lines that we investigated. IV1 is a 25-mer oligonucleotide localized in intron 1, 216/217 bases upstream from the start of exon 2. The human and mouse are identical except that the human sequence contains an extra G nucleotide at position 996. IV1 primer corresponding to the mouse sequence was synthesized. The 23-mer EX2R ends 28 nucleotides upstream from the 3′ end of exon 2. EX5 is a 22-mer, beginning 34 bases downstream from the 5′ end of exon 5. NCOI represents a 20-bp, 100% conserved, DNA fragment, starting 112/113 bases downstream from the TGA stop codon and contains the unique NCOI site present in the human and mouse gene.

An ATG primer was also synthesized, corresponding to a 20 oligonucleotide stretch, extending both 5′ and 3′ from the initiator methionine codon. This sequence is also totally conserved between primates and mouse. Whereas combinations of exonic primers (ATG/EX2R, EX5/NCOI, and ATG/NCOI) were able to direct amplification from cDNA prepared from RNA of the human Epo-producing cell line Hep3B, no amplification was obtained on cDNA prepared

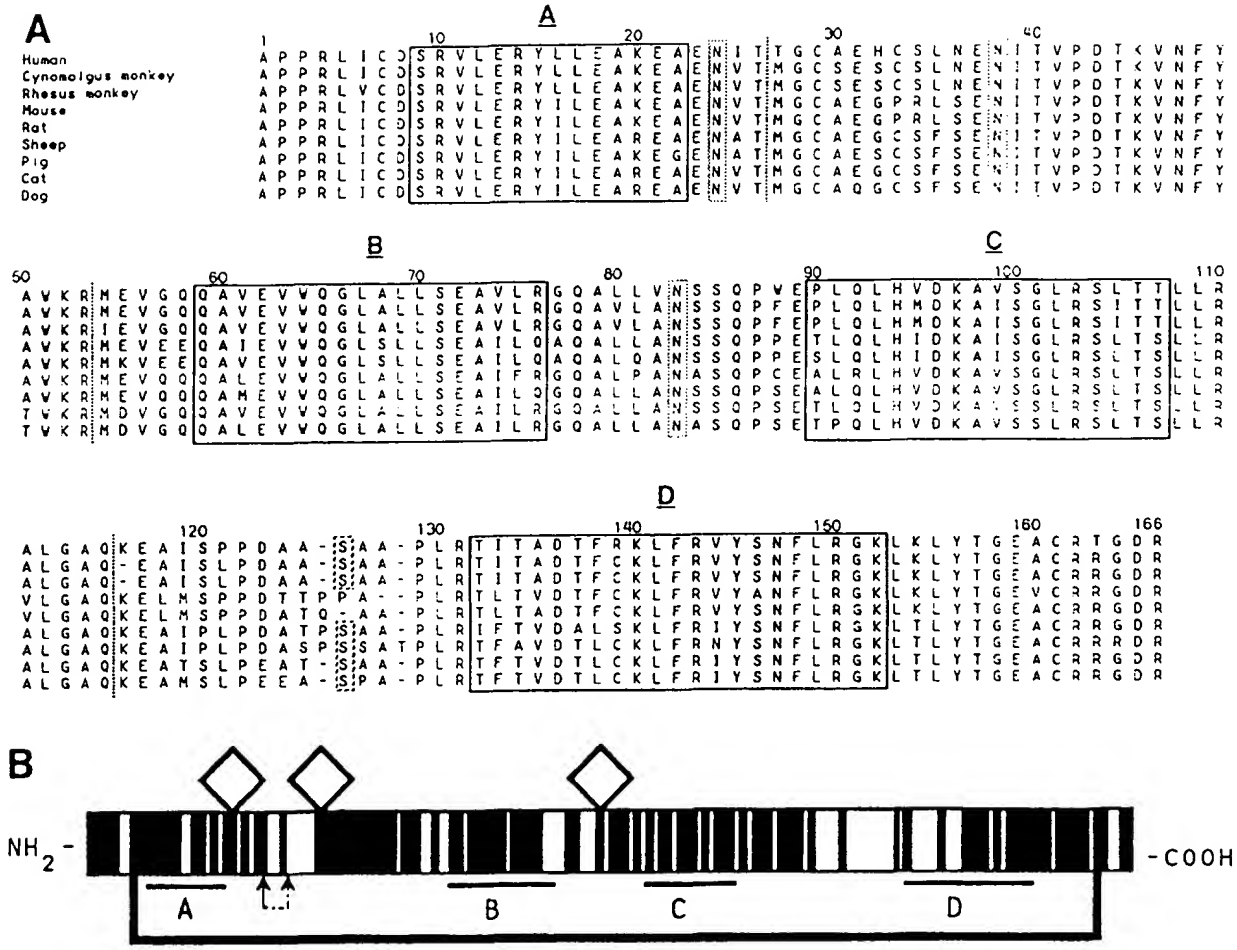


Fig 6. (A) Alignment of the primary structures of mature mammalian Epo proteins. The human, *Macaca fascicularis*, and mouse amino acid sequences were previously reported.[15-17] The residue numbers correspond to the 166 amino acids of the mature human protein. Plain boxes indicate the positions of the predicted four α helices in the human sequence.[19] Dashed boxes show the N- and O-glycosylation sites. Boundaries between exons are indicated by the vertical lines. (B) Schematic representation of mammalian Epo. The invariant amino acids among the eight sequences are represented by the solid boxes. The localization of each of the four α helices is underlined. The three glycosylation sites are shown as diamonds. The main disulfide bridge is indicated. The small arrows under the sequence indicate the short SH-bridge, which is missing in the rodents.

from uninduced or hypoxia-induced BHK, MDOK, LLC-PK1, MDCK, CRF-K, PAL1-K, and SP1-K cells (results not shown). Therefore, these renal cell lines apparently do not express Epo mRNA, either constitutively or after hypoxic stress, at a level detectable even by RT-PCR. Furthermore, no Epo protein was detectable by radioimmunoassay in the supernatants of cells maintained overnight in 1% O_2, 5% CO_2, and 94% N_2 at 37°C.

The amplification of species-specific IV1/EX2R and EX5/NCOI genomic fragments allowed us to design a strategy shown in Fig 2 for obtaining cDNA clones containing the complete coding sequence of mature Epo protein from various mammals. The IV1/EX2R primer pair resulted in a genomic fragment (~330 bp) extending from the 3' third of the first intron through 80% of exon 2, encoding the COOH-end of the propeptide and the NH_2-terminal amino acids of the mature protein. EX5/NCOI primers amplified a segment of exon 5 containing the COOH-end of the Epo protein and a 3' untranslated sequence of about 114 bp downstream of the stop codon. DNA sequencing of these 5' and 3' fragments permitted, if necessary, the design of a second generation of species-specific primers (SP) localized outside the coding part of the mature protein (ie, in the sequences encoding the propeptide and the 3' untranslated region).

Amplification of partial intron 1/exon 2 genomic clones. We first explored the efficacy of the IV1 and EX2R primers for amplification of genomic fragments from different purified DNA templates. In all the reactions, fragments of predicted size were the main (and, when using the most stringent annealing temperature, often the only) products detected by agarose gel electrophoresis. The (direct) sequences of these PCR-generated fragments are presented in Fig 3. The nucleotide alignment includes 6 different orders of mammals: one primate, the human (*Homo sapiens*); two artiodactyls, the sheep (*Ovis aries*) and pig (*Sus scrofa*); one perissodactyl, the horse (*Equus caballus*); one cetacean, the spotted dolphin (*Stenella plagiodon*); three carnivores, the dog (*Canis familiaris*), cat (*Felis catus*), and lion (*Panthera*

Table 1. Degree of Conservation Among Mammals of the Mature Epo Proteins

	% Identity
Primates	
Human/Rhesus monkey	90.5
Rhesus/Cynomolgus	98.8
Rodents	
Human/mouse	79.8
Human/rat	82.1
Mouse/rat	93.5
Artiodactyls	
Human/sheep	81.0
Human/pig	81.6
Sheep/pig	88.7
Carnivores	
Human/cat	84.5
Human/dog	84.5
Cat/dog	95.2

The percentages of identity among the various sequences were determined from the amino acid alignment reported in Fig 6A.

leo); and two rodents, the mouse (*Mus musculus*) and hamster (*Cricetus cricetus*).

The cat and lion exhibited an almost complete sequence identity, with only one T to G nucleotide substitution near the 5' end of the amplified portion of the first intron.

The ability of the intronic IV1 to anneal to genomic DNA from various species and, in combination with EX2R, to amplify PCR fragments of identical size, showed a high degree of conservation of sequence and position between mammals. This finding suggests that IV1 sequence may be involved to some extent in the regulation of the Epo gene expression. The comparison of IV1/EX2R sequences also showed two remarkably conserved intronic sequences: AT(T/A)GAATGAA(G/C)GC (nucleotides 1046 to 1058 in the human gene) and A(A/T)GGTN(G/C)GGG (nucleotides 1085 to 1094).

Amplification of partial exon 5 fragments from genomic DNAs. PCR reactions were also performed applying the EX5/NCOI primer pair to various purified genomic DNAs. As we previously observed for the amplification of IV1/EX2R, in all of the tested samples a unique main band of about 250 bp (as predicted for human and mouse genes) was detected on analytical agarose gel electrophoresis. Direct DNA sequencing of these fragments showed that the bulk of the purified PCR product corresponded to the expected Epo exon 5 sequence. However, for some DNA samples, which were purified from horse kidney and several cell lines (BHK and LLC-PK1 in particular), unknown sequences were present to various extents. These minor contaminating PCR products generated extraneous peaks in the sequence data, resulting in ambiguities at several positions. Nevertheless, computer-edited analyses of generated 3' noncoding sequences were sufficient to design, if necessary, specific primers with greater than 95% accuracy (data not shown).

Cloning of partial cDNAs encoding the complete mature Epo protein. As we previously mentioned, we were unable to obtain any amplification with cDNAs prepared from renal-derived mammalian cell lines. Therefore, we tested our battery of primers on reverse-transcribed RNAs prepared from kidney of several mammals. We have successfully amplified Epo cDNA from one primate, the Rhesus monkey (*Macaca mullata*); two carnivores, the dog (*Canis familiaris*) and cat (*Felis catus*); one rodent, the rat (*Rattus norvegicus*); and two artiodactyls, the pig (*Sus scrofa*) and sheep (*Ovis aries*). The aligned nucleotide sequences are presented in Fig 4. They have been deposited with the GenBank Data Library and were given the following accession numbers: L10609 (Rhesus monkey), L10606 (cat), L10608 (rat), L10607 (pig), L10610 (sheep), and L13027 (dog).

The Rhesus monkey, rat, and sheep cDNAs were amplified using the ATG/NCOI primer-pair combination. The pig, cat, and dog sequences were obtained by the use of a 5' 24-mer specific primer (SP1, 5' TGCTTCTGCTATCTTTGCTGCTGC 3'). IV1/EX2R sequencing showed that this exon 2 sequence was shared by the three species. In all PCR amplifications, NCOI was used as the reverse primer. Conservation of the NCOI sequence among mammals suggests a potential biologic role of this nucleotide sequence (perhaps in modulating the stability of Epo mRNA).[29] Considerable homology is also found in the sequenced 3' untranslated segment of exon 5 (Fig 4).

When compared with human Epo, the overall percentages of nucleotide identity are, respectively, 93.5% for the Rhesus monkey, 86.7% for the cat, 85.7% for the dog and the pig, 83% for the sheep, and 75.5% for the rat. The Rhesus monkey and Cynomologus monkey[16] were 99.6% identical. The mouse[17,18] and rat nucleotide sequences showed greater than 93% identity. The cat and dog sequences were 91% identical and the two sequenced artiodactyls, the pig and sheep, showed 88% identity.*

Comparison of Mammalian Epo Primary Sequences

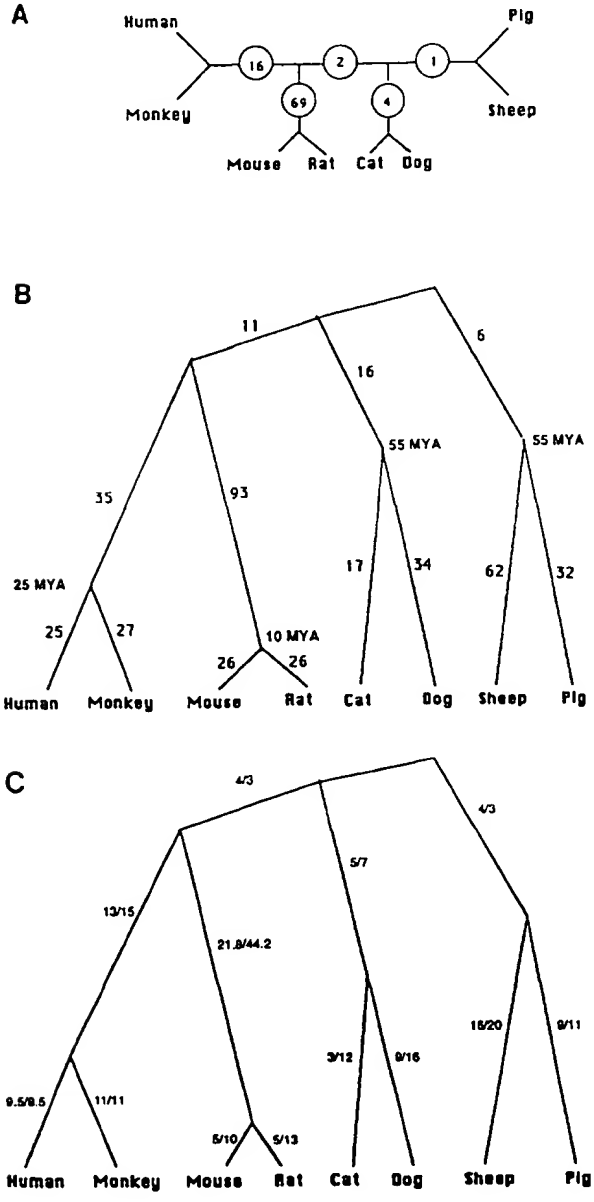
Propeptides. The predicted amino acid propeptide sequences of several mammals are presented in Fig 5. In the primates, there is strong conservation of the sequences of deduced propeptides, with only two amino acid substitutions at positions −11 (Leu in humans *v* Val in both monkeys) and −2 (Leu in humans *v* Pro in both monkeys). Expression of Cynomolgus monkey Epo gene in cultured mammalian cells resulted in the production of mature monkey Epo that was elongated at the N-terminus by three additional residues, Val-Pro-Gly.[16] As the Rhesus monkey has the same substitutions, an identical site of cleavage is likely. The Leu to Pro amino acid replacement at −2 is probably responsible for the differential activity of the signal peptidase observed between humans and monkeys. The three rodents (mouse, rat, and hamster) have identical propeptide amino acid sequences. The two carnivores also share one propeptide sequence.

Mature protein. A comparison of the three previously published and our six deduced amino acid sequences of mature Epo proteins is shown in Fig 6A. Epo is highly con-

* Several PCR attempts on total RNA purified from the horse kidney were unsuccessful, even when using 5- and 3-specific equine primers (SP), presumably because analysis on 1% formaldehyde/agarose gel showed degradation of the RNA.

served among mammals. The amino acid alignment showed that more than 63% of the molecule is composed of invariant amino acids (106 residues) (Fig 6B). Most of the observed substitutions are conservative, involving residues with similar physical and chemical properties. The calculated percentages of sequence identity are shown in Table 1.

There is 100% conservation of Cys7 at the N-terminal portion of the polypeptide and Cys161 near the C-terminus. This disulphide bridge is essential to the formation of a stable and functional cytokine.[30] In contrast, whereas Cys29 is invariant in all the species we examined, Cys33 is present only in the primates, sheep, pig, and cat, but not in the two rodents, in which there is a proline. The lack of functional importance of this short disulfide loop is underscored by muteins, reported by Boissel et al,[19] in which tyrosine substitutions at each or both of these sites have no effect on Epo's biologic activity.[19]



There is also 100% conservation of all three asparagine residues responsible for N-linked glycosylation. Even though Epo deprived of its carbohydrate either by enzymatic cleavage[31] or by production in bacteria[32] has full in vitro biologic activity, survival of the hormone in the circulation depends on N-linked glycosylation.[33] In contrast, the O-linked glycosylation site (Ser126 in human Epo) is not essential for biologic function as it is missing in mouse and rat Epos.

As shown in Fig 6A, there is very high conservation of sequence in regions that, by algorithms based on primary structure, are predicted to be α helices.[19] The few differences in sequence within these helical regions are conservative replacements. Indeed, some sites within these predicted α helices are conserved in corresponding helices of other cytokines.[34] In contrast, regions predicted to be interhelical loops are less well conserved, particularly the 14 residue stretch between residues 116 and 130 of human Epo. In a companion report,[19] we show that a mutein with a deletion of residues 111 through 119 has normal stability and biologic activity, whereas a deletion of residues 122 through 126 fails to produce a detectable protein, probably owing to markedly impaired stability.

Similar overall 4 α-helical bundle structure exists or is predicted for several other cytokines/hormones, such as growth hormone (GH), granulocyte-macrophage colony-stimulating factor (GM-CSF), interleukin-3 (IL-3), IL-4, and IL-5.[35] Like Epo, GH exhibits a high degree of primary sequence conservation and biologic cross-reactivity between mammals. The mouse, rat, pig, and sheep proteins show about 80% of amino acid identity with human GH. On the contrary, in other cytokines there is more sequence diversity among species. For example, the human and murine forms of IL-5, GM-CSF, and IL-3 show, respectively, only 69%, 54%, and 26% amino acid sequence identity. Fur-

Fig 7. Phylogenetic lineages derived from analyses of mammalian Epo cDNA sequences encoding the full-length mature proteins from seven species. (A) Strength of groups in the maximum parsimony tree found on examining all 945 unrooted trees formed by seven terminal taxa. This tree of lowest length required 374 base substitutions. Each link between ancestral nodes has a circled number; this strength of grouping number is the minimum number of substitutions that must be added to the length of the maximum parsimony tree to find a tree that breaks down the barrier (moves one or more sequences) between the two groups separated by the interior link. (B) The phylogenetic tree derived from the maximum parsimony reconstruction. On the basis of other molecular evidence involving comparative amino acid sequence data from monotremes, marsupials, and many eutherian species,[36,38] the root of this Epo phylogenetic tree is placed on the interior link separating the artiodactyl group from the primate, rodent, and cat group. The numbers on the internodal links represent the numbers of base pairs by which the nodal ancestral and descendant sequences differ. MYA, millions of years ago, as inferred from paleontologic views of eutherian phylogeny. (C) Parsimony reconstruction on that portion of the cDNA sequences that are codons for amino acids. The numbers shown as a fraction on each link are, in the numerator, the number of amino acid changing base replacements and, in the denominator, the number of silent base replacements. The computer algorithm that performed this calculation is described in Czelusniak et al[42] (see the Fig 2 legend of this reference).

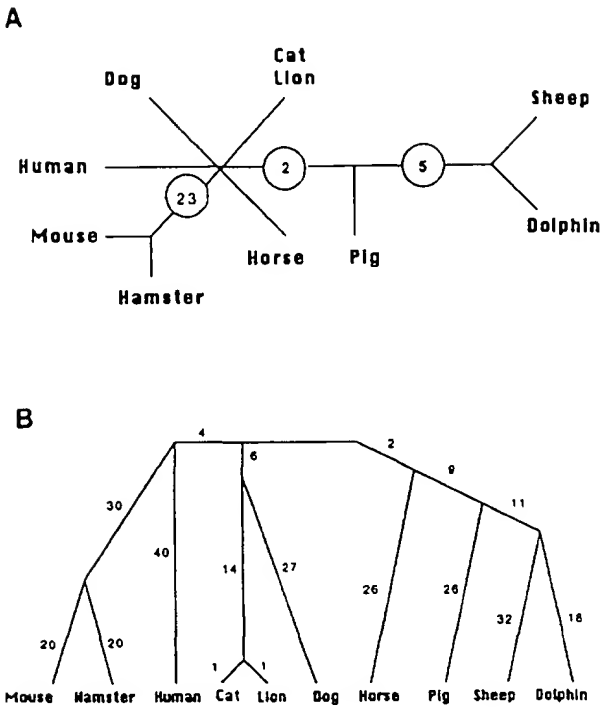


Fig 8. Phylogenetic lineages derived from analyses of the mammalian Epo intron 1-exon 2 sequences. (A) Strength of groups in the maximum parsimony tree found on examining all 135,135 trees formed by nine terminal taxa. As the two feloid sequences (those of cat and lion) are nearly identical, they were treated as a single taxon in this strength of grouping analysis. The maximum parsimony tree for the segment (about 285 bp) of intron 1 and exon 2 from 10 species required 286 base substitutions. The circled numbers are the strength of grouping numbers (as defined in the Fig 7A legend). (B) The near most parsimonious tree that groups horse with the artiodactyl-cetacean branch. This tree required 287 base substitutions. The numbers on links represent numbers of base pairs by which the nodal ancestral and descendant sequences differ.

thermore, unlike Epo or GH, those proteins lack interspecies biologic cross-reactivity.

Expression of the mammalian Epo hormones. Mammalian Epo cDNAs were subcloned into the pSG5 plasmid and transiently expressed in the monkey Cos7 cell line. Supernatants of the transfected cells were able to sustain the cellular proliferation of the murine Epo-dependent HCD57 cell line, showing the biologic cross-reactivity between species. As the human Epo antisera used in our radioimmunoassay bind to Epos from other species with variable affinity, the amount of produced protein was difficult to determine accurately. Experiments are in progress to further characterize the relative affinities of the various mammalian Epos toward murine and human Epo receptors.

Phylogenetic analyses of Epo sequences. Nucleotide sequences encoding the full-length mature Epo protein were analyzed by the maximum parsimony method.[36,37] An all trees set of computer programs determined the parsimony length of each of the 945 unrooted trees formed by the seven cDNA sequences (the terminal taxa) (Fig 4) and, on ordering those 945 trees according to increasing length, identified the minimum number of extra nucleotide substitutions needed to break up each group in the maximum parsimony (lowest length) tree. As shown in Fig 7A and B, the mouse and rat are strongly grouped, as are the human and monkey. The dog and cat are not as strongly grouped, and the pig and sheep are very weakly grouped. These grouping results are in accord with accepted phylogenetic views, which place the intraordinal divergence, on the one hand, between dog and cat and, on the other hand, between pig and sheep at about 50 to 60 million years ago. As expected with the parsimony reconstruction on only the codon sequences, silent base substitutions occurred much more readily than did amino acid replacements (Fig 7C).

The Epo intron 1-exon 2 sequences shown in Fig 3 were also analyzed by the maximum parsimony method. Figure 8A shows strength of grouping results based on all 135,135 unrooted trees formed by nine terminal taxa (in this case, the cat and lion sequences, which are nearly identical, were treated as a single taxon). Because of the quite limited phylogenetic information in these relatively short sequences, there are alternative maximum parsimony trees (three were found) and the consensus of these alternative trees does not depict (Fig 8A) a consistently bifurcating tree. Nevertheless, the rodents (mouse and hamster) are strongly grouped, as are the feloids (cat and lion); at more moderate strength, the cetacean (dolphin) is grouped with the artiodactyls (in particular, with sheep). There are other molecular data as well as some paleontologic data that depict cetaceans originating from early artiodactyls.[36,38-41] Figure 8B shows a near maximum parsimony tree. This tree has a length of 287 base substitutions, whereas the length of each maximum parsimony tree is 286 base substitutions. A maximum parsimony tree can be obtained by exchanging positions of the horse and human branches, thus placing the human rather than the horse close to artiodactyls. However, the grouping of horse with the artiodactyl-cetacean clade (ie, the near maximum parsimony tree) is in better agreement with well-supported phylogenetic evidence on mammalian ordinal relationships than is the maximum parsimony tree.

ACKNOWLEDGMENT

We thank Hanna Bandes for her able secretarial and editorial assistance.

REFERENCES

1. Jelkmann W: Erythropoietin: Structure, control of production, and function. Physiol Rev 72:449, 1992

2. Semenza GL, Nejfelt MK, Chi SM, Antonarakis SE: Hypoxic-inducible nuclear factors bind to an enhancer element located 3' to the human erythropoietin gene. Proc Natl Acad Sci USA 88:5680, 1988

3. Beck I, Ramirez S, Weinmann R, Caro J: Enhancer element at the 3'-flanking region controls transcriptional response to hypoxia in the human erythropoietin gene. J Biol Chem 266:15563, 1991

4. Pugh CW, Tan CC, Jones RW, Ratcliffe PJ: Functional analysis of an oxygen-regulated transcriptional enhancer lying 3' to the mouse erythropoietin gene. Proc Natl Acad Sci USA 88:10553, 1991

5. Blanchard KL, Acquaviva AM, Galson DL, Bunn HF: Hypoxic induction of the human erythropoietin gene: Cooperation be-

tween the promoter and enhancer, each containing steroid receptor response elements. Mol Cell Biol 12:5373, 1992

6. Garcia JF, Schooley JD: Immunological neutralization of various erythropoietins. Proc Soc Exp Biol Med 112:712, 1963

7. Zanjani ED, Gordon AS, Wong KK, McLaurin WD: The renal erythropoietic factor (REF) X. The question of species and class specificity. Proc Soc Exp Biol Med 131:1095, 1969

8. Biljanovic-Paunovic L, Basara N, Milenkovic P, Pavlovic-Kentera V: Mammalian erythropoietin species specificity detected in vivo and in vitro. Period Biol 90:421, 1988

9. Rosse WF, Waldmann T, Hull E: Factors stimulating erythropoiesis in frogs. Blood 22:66, 1963

10. Pinder A, Burggren WJ: Respiration during chronic hypoxia and hyperoxia in larval and adult bullfrogs (rana catesbeiana): Changes in respiratory properties of whole blood. Exp Biol 105:205, 1983

11. Rosse WF, Waldmann TA: Factors controlling erythropoiesis in birds. Blood 27:654, 1966

12. Black CP, Tenney SM: Oxygen transport during progressive hypoxia in high-altitude and sea-level waterfowl. Respir Physiol 39:217, 1980

13. Zanjani ED, Yu M-L, Perlmutter A, Gordon AS: Humoral factors influencing erythropoiesis in the fish (Blue Gourami—*Trichogaster trichopterus*). Blood 33:573, 1969

14. Jacobs K, Shoemaker C, Rudersdorf R, Neill SD, Kaufman RJ, Mufson A, Seehra J, Jones SS, Hewick R, Fritsch EF, Kawakita M, Shimizu T, Miyake T: Isolation and characterization of genomic and cDNA clones of human erythropoietin. Nature 313:806, 1985

15. Lin F-K, Suggs S, Lin C-H: Cloning and expression of the human erythropoietin gene. Proc Natl Acad Sci USA 82:7580, 1985

16. Lin F-K, Lin C-H, Lai P-H, Browne JK, Egrie JC, Smalling R, Fox GM, Chen KK, Castro M, Suggs S: Monkey erythropoietin gene: Cloning, expression and comparison with the human erythropoietin gene. Gene 44:201, 1986

17. McDonald JD, Lin KF, Goldwasser E: Cloning, sequencing, and evolutionary analysis of the mouse erythropoietin gene. Mol Cell Biol 6:842, 1986

18. Shoemaker CB, Mitsock LD: Murine erythropoietin gene: Cloning, expression, and human gene homology. Mol Cell Biol 6:849, 1986

19. Boissel J-P, Lee W-R, Presnell SR, Cohen F, Bunn HF: Erythropoietin structure-function relationships: Mutant proteins that test a model of tertiary structure. J Biol Chem (in press)

20. Blin N, Stafford DW: A general method for isolation of high molecular weight DNA from eukaryotes. Nucleic Acids Res 3:2303, 1976

21. Chirgwin JM, Pryzbyla AE, MacDonald RJ, Rutter WJ: Isolation of biologically active ribonucleic acid from sources enriched with ribonuclease. Biochemistry 18:5294, 1979

22. Tracy TE, Mulcahy LS: A simple method for direct automated sequencing of PCR fragments. Biotechniques 11:68, 1991

23. Nagao M, Suga H, Okano M, Masuda S, Narita H, Ikura K, Sasaki R: Nucleotide sequence of rat erythropoietin. Biochim Biophys Acta 1171:99, 1992

24. Sawyer ST, Hankins WD: Erythropoietin receptor metabolism in erythropoietin-dependent cell lines. Blood 72:132a, 1988 (abstr, suppl 1)

25. Beru N, McDonald J, Lacombe C, Goldwasser E: Expression of the erythropoietin gene. Mol Cell Biol 6:2571, 1986

26. Schuster SJ, Wilson JH, Erslev AJ, Caro J: Physiologic regulation and tissue localization of renal erythropoietin messenger RNA. Blood 70:318, 1987

27. Bondurant MC, Koury MJ: Anemia induces accumulation of erythropoietin mRNA in the kidney and liver. Mol Cell Biol 6:2731, 1986

28. Fandrey J, Bunn HF: In vivo and in vitro regulation of erythropoietin mRNA: Measurement by competitive polymerase chain reaction. Blood 81:617, 1993

29. Rondon IJ, MacMillan LA, Beckman BS, Goldberg MA, Schneider T, Bunn HF, Malter JS: Hypoxia up-regulates the activity of a novel erythropoietin mRNA binding protein. J Biol Chem 266:16594, 1991

30. Wang FF, Kung CK-H, Goldwasser E: Some chemical properties of human erythropoietin. Endocrinology 116:2286, 1985

31. Takeuchi M, Takasaki S, Shimada M, Kobota A: Role of sugar chains in the in vitro biological activity of human erythropoietin produced in recombinant Chinese hamster ovary cells. J Biol Chem 265:12127, 1990

32. Narhi LO, Arakawa T, Aoki KH, Elmore R, Rohde MF, Boone T, Strickland TW: The effect of carbohydrate on the structure and stability of erythropoietin. J Biol Chem 266:23022, 1991

33. Spivack JL, Hogans BB: The in vivo metabolism of recombinant erythropoietin in the rat. Blood 73:90, 1989

34. Manavalan P, Swope DL, Withy RM: Sequence and structural relationships in the cytokine family. J Protein Chem 11:321, 1992

35. Bazan JF: Haemopoietic receptors and helical cytokines. Immunol Today 11:350, 1990

36. Czelusniak J, Goodman M, Moncrief ND, Kehoe SM: Maximum parsimony approach to construction of evolutionary trees from aligned homologous sequences. Methods Enzymol 183:601, 1990

37. Stanhope MJ, Czelusniak J, Si J-S, Nickerson J, Goodman M: A molecular perspective on mammalian evolution from the gene encoding interphotoreceptor retinoid binding protein, with convincing evidence for bat monophyly. Molec Phylogenet Evol 1:148, 1992

38. Czelusniak J, Goodman M, Koop BF, Tagre DA, Shoshani J, Braunitzer G, Kleinschjidt TK, DeJong WW, Matsuda G: Perspectives from amino acid and nucleotide sequences on cladistic relationships among higher taxa of eutheria, in Genoways HH (ed): Current Mammology, vol 2. New York, NY, Plenum, 1990, p 545

39. Irwin DM, Kocher TD, Wilson AC: Evolution of the cytochrome b gene of mammals. J Mol Evol 32:128, 1991

40. Gingerich PD, Smith BH, Simons E: Hind limbs of eocene basilosaurus: Evidence of feet in whales. Science 249:154, 1990

41. Novacek MJ: Mammalian phylogeny: Shaking the tree. Nature 356:121, 1992

42. Czelusniak J, Goodman M, Hewett-Emmett D, Weiss ML, Venta PJ, Tashian RE: Phylogenetic origins and adaptive evolution of avian and mammalian haemoglobin genes. Nature 298:297, 1982

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 269, No. 36, Issue of September 9, pp. 22839-22846, 1994
Printed in U.S.A.

Erythropoietin Structure-Function Relationships

IDENTIFICATION OF FUNCTIONALLY IMPORTANT DOMAINS*

(Received for publication, April 22, 1994, and in revised form, June 10, 1994)

Danyi Wen‡, Jean-Paul Boissel‡, Mark Showers‡, Baird C. Ruch, and H. Franklin Bunn§

From the Hematology/Oncology Division, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115

In order to delineate functionally important domains in erythropoietin (Epo), we have prepared and tested a series of amino acid replacements at 51 conserved sites predicted to be on the surface of the molecule. Alanine replacements permitted preservation of α-helical structure. Wild type and mutant Epo cDNAs were transiently expressed at high levels in COS1 and COS7 cells. The biological activity of wild type and mutant Epos was assayed in three Epo-responsive cell types: primary murine erythroid spleen cells, the murine HCD57 erythroleukemia cell line, and the human UT7-EPO leukemia cell line. When Arg^{14} on predicted Helix A was replaced by Ala, biological activity was substantially reduced, whereas replacement with Glu resulted in total loss of specific bioactivity. In a similar manner, the mutein $Arg^{103} \rightarrow$ Ala in Helix C was completely lacking in biological activity, whereas both $Ser^{104} \rightarrow$ Ala and $Leu^{106} \rightarrow$ Ala had decreased bioactivity. In Helix D, the mutein $Gly^{151} \rightarrow$ Ala had markedly decreased bioactivity, whereas that of the adjacent $Lys^{152} \rightarrow$ Ala mutein was moderately impaired. In contrast, Ala replacements at three nearby sites on Helix D (147, 146, and 143) resulted in muteins with increased bioactivity. In conclusion, our mutagenesis experiments have identified functionally important domains on the surface of the Epo molecule, at sites comparable with those established for other cytokines.

Erythropoietin (Epo)[1] is a 30.4-kDa glycoprotein hormone that regulates red blood cell production (1, 2). Biochemical studies (3, 4) led to molecular cloning of the Epo gene (5, 6). High level expression has enabled recombinant human Epo to be used very effectively in the treatment of anemias. Moreover, the experimental use of recombinant Epo has greatly advanced our understanding of the molecular mechanisms underlying erythropoiesis.

Epo is a member of an extensive cytokine family which also includes growth hormone, prolactin, interleukins 2 through 7, G-CSF, GM-CSF, M-CSF, oncostatin M, leukemia inhibitory factor, and ciliary neurotrophic factor (7–9). Although sequence homology is weak, genes encoding these proteins have similar numbers of exons and share a clear relationship between intron-exon boundaries and secondary structure (9). All of these cytokines are predicted to fold into a common compact globular structure, consisting of four amphipathic α-helical bundles. Such theoretical models of the structures of human growth hormone (HGH) (10) and IL-4 (11) have been in good agreement with the subsequent structures established by x-ray diffraction (HGH) (12, 13) or by multidimensional NMR (IL-4) (14, 15). Moreover, the crystal structures of GM-CSF (16, 17), G-CSF (18), monomeric M-CSF (19), IL-2 (20, 21), and IL-5 (22) are also in reasonable agreement with their predicted structures (for review, see Ref. 23).

Epo's high carbohydrate content (39%) and the heterogeneity of these carbohydrate chains have thus far precluded the determination of its three-dimensional structure by either x-ray diffraction or by NMR. In order to gain an understanding of Epo structure-function relationships, we first determined the coding sequences of Epo from six mammals representing five different orders, thereby identifying regions of high homology that are likely to have structural and/or functional importance (24). We then developed a computer-generated model of Epo's three-dimensional structure and have tested it by means of a series of scanning deletion muteins, which fully support a four α-helical bundle structure (8, 25) in common with that of the other cytokines mentioned above. As described in this report, we have exploited this information in the design and testing of single amino acid replacement muteins that provide information on functionally important domains of human Epo.

MATERIALS AND METHODS

Construction of Wild Type and Mutant Mammalian Expression Plasmids—λHEPO FL12, kindly provided by Genetics Institute (Cambridge, MA) (5), is an M13 plasmid, containing a 1.4-kilobase pair *Eco*RI-*Eco*RI human Epo cDNA insert. A 943-base pair *Eco*RI-*Bgl*II fragment, which includes the complete coding sequence of the wild type human erythropoietin as well as untranslated regions 216 base pairs upstream and 183 base pairs downstream, was inserted into the mammalian expression plasmid pSG5 (Stratagene) (26) and designated pSG5-EPO/WT.

Site-directed Mutagenesis—Site-directed mutagenesis was carried out by a modification of the protocol described by Kunkel *et al.* (27). For further details see Ref. 25. Since a 40–80% mutation yield is normally obtained, four to five double-stranded cDNA clones from each reaction were sequenced with 7-deaza-dGTP and Sequenase (U. S. Biochemical Corp.) (28) to verify the mutation. In addition, each mutant was screened by restriction mapping or, less commonly, by full sequencing in order to detect the presence of additional unwanted mutations.

Production of Wild Type and Epo Muteins in Mammalian Cells—COS 1 or COS7 cells, grown to 40–60% confluence, were transfected with 20 μg of recombinant plasmid DNA/10-cm dish using the calcium phosphate precipitation protocol (29). As a control of transfection efficiency, in several experiments 2 μg of pCH110 plasmid (Pharmacia Biotech Inc.) was co-transfected and β-galactosidase activity was measured in the cytoplasmic extracts.

Quantitation of Transiently Expressed Recombinant Epos—The amount of secreted protein in the supernatants of transfected COS7

* This work was supported by National Institutes of Health Grant RO1-HL42949 and by a grant from the R. W. Johnson Pharmaceutical Research Institute. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Drs. Boissel, Showers, and Wen contributed equally to this work and should all be regarded as first authors.

§ To whom correspondence should be addressed: Hematology/Oncology Division, Brigham and Women's Hospital, 221 Longwood Ave., LMRC-2, Boston, MA 02115. Tel.: 617-732-5841; Fax: 617-739-0748.

[1] The abbreviations used are: Epo, erythropoietin; CSF, colony-stimulating factor; G-CSF, granulocyte CSF; M-CSF, macrophage CSF; GM-CSF, granulocyte-macrophage CSF; HGH, human growth hormone; GH, growth hormone; IL, interleukin.

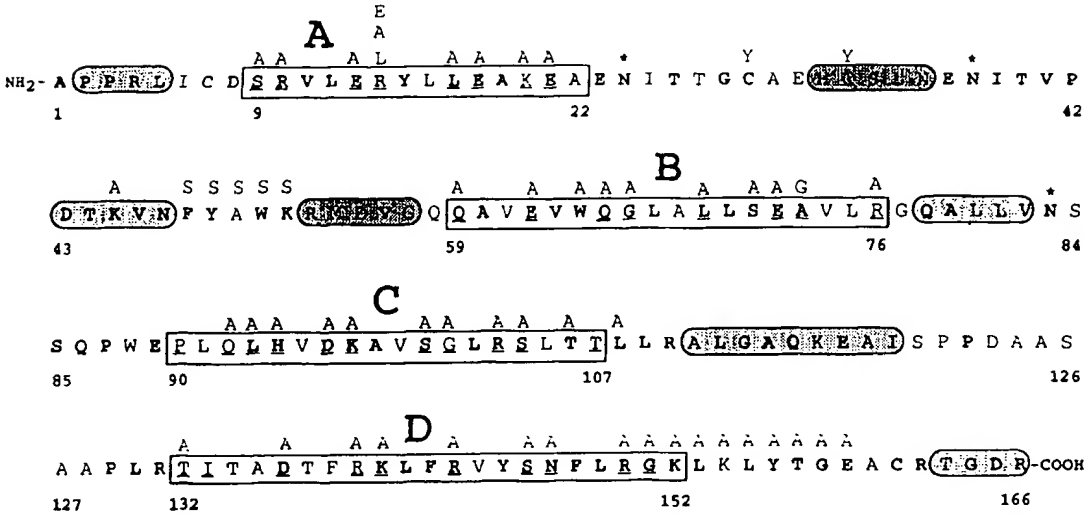


FIG. 1. **Primary and secondary structure of human Epo.** The amino acid sequence of the mature exported protein is shown by the single letter code. Sequence letters in *bold* are fully conserved among eight mammalian species (24). Each of the three *N*-linked glycosylation sites is designated by an *asterisk*. The predicted amphipathic α helices (A, B, C, D) are enclosed in *rectangular boxes*. Residues predicted to be on the surface are *underlined*. The *shaded patches* designate residues in the nonhelical regions of human Epo which can be deleted with no significant loss of specific bioactivity (25). The amino acid replacement muteins reported in this study are shown above the primary sequence.

was determined by a radioimmunoassay which utilized a high titer rabbit polyclonal antiserum raised against human wild type Epo and produced in our laboratory (30). This polyclonal antiserum is expected to measure the concentration of Ala replacement Epo muteins more accurately than monoclonal antisera. ^{125}I-Labeled recombinant human Epo was obtained from Amersham Corp. (for further details, see Ref. 31). In order to obtain accurate Epo concentrations, several dilutions of each COS cell supernatant were assayed in duplicate.

Bioassays—The dose-dependent proliferation activities of wild type Epo and Epo muteins were assayed *in vitro* using three different Epo-responsive target cells: 1) primary murine spleen cells, following a modification of the method of Krystal (32); 2) HCD 57, a murine MEL cell line developed by Hankins *et al.* (33); and 3) UT7-Epo, a human cell line derived from the bone marrow of a patient with acute megakaryoblastic leukemia (34). After 22–72 h of incubation (depending on cell type) with increasing amounts of recombinant Epo proteins, cellular growth was determined by [^{3}H]thymidine (New England Nuclear) incorporation (35). For each assay, nine concentrations of Epo were used (2.5–50 milliunits/ml). The counts/min in the absence of any source of Epo (nonspecific background) was subtracted from each experimental point. In order to maximize the accuracy of these bioassays, the following controls were employed. (*a*) In each experiment COS cell supernatants containing Epo muteins were compared with COS cell supernatants from three separate transfections of wild type Epo. (*b*) To correct for any nonspecific effect of other substances in the COS cell supernatant on the growth of cells in the bioassay, wild type recombinant human Epo (rhEpo) was added to mock-transfected COS cells and to media alone and assayed for bioactivity. These specific bioactivity values were generally identical to those of wild type rhEpo secreted from the COS cells. (*c*) For muteins with abnormal specific bioactivity, replicate transfections were done in order to minimize possible variability in the contribution of nonspecific factors in the COS cell supernatants in the measurements of biological activity.

RESULTS

Our objective is to identify the functionally important domains of the Epo molecule by preparing muteins that have altered specific bioactivity, *i.e.* significantly higher or lower biological activity per unit mass of protein. Such muteins must satisfy two criteria: 1) efficient secretion by COS7 cells and 2) preservation of the overall folding and three-dimensional structure, thereby allowing full recognition (with normal or nearly normal binding affinity) by the polyclonal anti-Epo antiserum that we use in our radioimmunoassay.

In order to satisfy these criteria, our strategy employed the use of muteins that have only single amino acid replacements. For all but a few muteins, we have made alanine substitutions in order to preserve α-helical structure (36). A comprehensive alanine replacement scan of Epo would require the preparation and testing of nearly 166 muteins. Such an undertaking would be prohibitively labor intensive and costly. We therefore devised a strategy to reduce the number of required replacements, by applying the information we recently obtained and reported on mammalian Epo sequences (24), as well as on scanning deletion muteins (25).

On the basis of its primary amino acid sequence and disulfide bonds, Epo is predicted to have a four-antiparallel amphipathic α-helical bundle structure (8, 25) in common with other members of the cytokine family (9). We have shown that deletions in nonhelical regions at the N terminus, the C terminus, and in the loops between helices result in the formation of Epo proteins that are readily secreted from the cell and are biologically active (25). These regions (Fig. 1) can be ruled out as domains essential for function, such as the sites involved in the binding of Epo to its receptor. Moreover, residues that are not conserved among mammals (Ref. 24 and Fig. 1) are also unlikely to be functionally important sites. Our previous scanning deletion analysis (25) suggests that the functionally important receptor contact sites are likely to reside on the predicted α helices. Furthermore, since the amphipathic helices are predicted to bind to one another along their hydrophobic surfaces, the biologically relevant contact sites are likely to be residues predicted to be on the external surfaces of the helices (25). These considerations greatly restricted the number of mutations needed to define critical functional domains of Epo and focused our analysis on the conserved surface residues of the four helices as well as residues adjoining the helices. In addition to the muteins which were selected according to the above criteria, we prepared 3 alanine replacement muteins at conserved sites predicted to be in the hydrophobic interior: Trp^{64}, Ser^{71}, and Thr^{106}. As expected (and shown in Table I), all of these muteins had normal specific bioactivity.

Fig. 1 depicts the residues excluded on the basis of the three above-mentioned criteria as well as the 56 replacement muteins we have prepared and tested.[2] Table I presents specific bioactivity data on all of these muteins tested in three Epo-responsive cell types. In general there was excellent agreement

[2] 53 muteins at 51 predicted external sites.

TABLE I
Bioactivity of Epo replacement muteins

The letters in the column on the left designate predicted helices (A, B, C, D) or interhelical loops (A-B, C-D). The amino acid of wild type human Epo, its position from the N terminus, and its replacement according to the single amino acid letter code are listed under the mutein column. The three bioassay columns (human UT7 cell line, murine HCD57 cell line, and primary mouse spleen erythroid cells) show specific bioactivity of each mutein, expressed as a percentage of wild type human Epo bioactivity, with the background COS supernatant alone subtracted from the value. The mean and standard deviation of the bioassay values are listed for $n \geq 3$ determinations. The numbers within the brackets indicate the number of separate COS cell transfections over (/) the number of separate bioassays that were performed. NS, not secreted from COS 7 cells.

Helix	Mutein	Specific bioactivity		
		UT7	HCD	Spleen
		% wild type human Epo		
A	S9A	148 ± 35; [3/3]	123 ± 13; [3/3]	
A	R10A	91; [1/1]	123; [1/1]	
A	E13A	111, 74; [2/2]	80, 99; [2/2]	
A	R14A	58, 56; [2/2]	2, 30; [2/2]	18, 24; [2/2]
A	R14L	91 ± 8; [4/6]	86 ± 15; [4/6]	67, 111; [2/2]
A	R14E	11 ± 6 [3/5]	12 ± 1; [5/3]	17, 17; [2/2]
A	L17A	110; [1/1]	95; [1/1]	
A	E18A	70 ± 17; [3/3]	100, 90; [2/2]	
A	K20A	252 ± 81; [4/4]	103 ± 21; [4/4]	
A	E21A	99, 69; [2/2]	50; [1/1]	
A-B	C29Y/C33Y	75; [1/1]	42; [1/1]	
A-B	K45A	110, 101; [2/2]	100, 100; [2/2]	
A-B	F48S	146, 110; [2/2]	100, 61; [2/2]	
A-B	Y49S	150 ± 7; [3/3]	146, 138; [2/2]	
A-B	A50S	128, 92; [2/2]	140, 37; [2/2]	
A-B	W51S	101, 78; [2/2]	89 ± 7; [3/3]	
A-B	K52S	99, 90; [2/2]	99, 88; [2/2]	
B	Q59A	156, 125; [1/2]	130, 134; [1/2]	
B	E62A	101, 85; [2/2]	63, 78; [2/2]	82, 111; [2/2]
B	W64A	102, 112; [2/2]	87, 143; [2/2]	90, 107; [2/2]
B	Q65A	96; [1/1]	140; [1/1]	99; [1/1]
B	G66A	130; [1/1]	103; [1/1]	85; [1/1]
B	L69A	94; [1/1]	84; [1/1]	
B	S71A	64; [1/1]	130; [1/1]	85; [1/1]
B	E72A	NS		
B	A73G	242 ± 25; [2/3]	104; [1/1]	
B	R76A	109; [1/1]	138; [1/1]	
C	Q92A	112; [1/1]	98, 95; [2/2]	113, 91; [2/2]
C	L93A	123, 126; [2/2]	127, 95; [2/2]	
C	H94A	NS		
C	D96A	NS		
C	K97A	84 ± 9; [2/4]	77 ± 27; [2/3]	94; [1/1]
C	S100A	98, 85; [2/2]	131; [1/1]	104; [1/1]
C	G101A	166 ± 52; [3/3]	146 ± 26; [3/4]	73, 152; [2/2]
C	R103A	3 ± 3; [3/5]	7 ± 1; [2/3]	6; [1/1]
C	S104A	114 ± 28; [2/4]	35 ± 17; [2/3]	
C	T106A	120; [1/1]	108; [1/1]	
C-D	L108A	158 ± 41; [2/4]	48 ± 6; [2/4]	12, 52; [2/2]
D	T132A	120, 104; [1/1]	98, 108; [1,1]	
D	D136A	108, 114; [2/2]	88 ± 20; [3/3]	
D	R139A	171, 91; [2/2]	120, 109; [2/2]	90, 131; [1/2]
D	K140A	189 ± 30; [3/3]	110 ± 19; [3/3]	
D	R143A	276 ± 85; [3/3]	150 ± 27; [3/3]	
D	S146A	198, 269; [1/2]	58, 110; [1/2]	
D	N147A	457 ± 185; [3/6]	115 ± 22; [5/6]	
D	R150A	181 ± 94; [3/4]	92 ± 28; [3/4]	28, 64; [1/2]
D	G151A	11 ± 4; [2/3]	6 ± 2; [2/3]	
D	K152A	89, 73; [1/2]	46, 20; [1/2]	16, 35; [1/2]
3'D	L153A	95; [1/1]	83; [1/1]	46; [1/2]
3'D	K154A	175 ± 60; [3/4]	185 ± 84; [3/5]	82, 157; [1/2]
3'D	L155A	95, 128; [2/2]	115; [1/1]	40, 98; [2/2]
3'D	Y156A	105; [1/1]	65 ± 12; [3/3]	
3'D	T157A	86, 104; [2/2]	89, 91; [2/2]	102, 104; [2/2]
3'D	G158A	84, 109; [2/2]	94, 89; [2/2]	77, 139; [2/2]
3'D	E159A	89, 97; [2/2]	101 ± 10; [2/3]	33, 96; [2/2]

between the three cell bioassay systems. As discussed below, a few muteins displayed different biological activity depending on the species of the assay cell, specifically the human UT7 *versus* the murine HCD57 and spleen cells.

It should be noted that in a few of these dose-response curves, at the highest concentrations of Epo tested, there was a drop-off in the incorporation of [^{3}H]thymidine (for example, see Figs. 4*B* and 5*C*). This is due to the fact that high levels of biologically active Epo cause terminal maturation and cessation of cell division. Because of this possibly confounding phenomenon, it is important to have, as in the experiments that we report here, a full dose-response curve for each mutein over a wide (50-fold) concentration range.

At three sites in human Epo, we found a marked decrease in specific bioactivity. (*a*) In predicted Helix A, replacement of Arg^{14} by Ala resulted in considerably impaired biological activity, particularly when assayed by the two mouse cell systems (Fig. 2). In order to further delineate residue 14 as a functionally important site, the following additional muteins were prepared and tested: $Arg^{14} \rightarrow$ Leu and $Arg^{14} \rightarrow$ Glu. As shown in Fig. 2, R14L had normal specific bioactivity, whereas R14E was nearly inactive, as determined by all three cell bioassays. The differential bioactivity of these Arg^{14} replacement muteins with alanine, leucine, or glutamic acid indicates a precise stereochemical integrity at this site that is essential for Epo's function. When other external sites predicted to be on Helix A were replaced by Ala (Ser^9, Arg^{10}, Glu^{13}, Leu^{17}, Glu^{18}, Lys^{20}, and Glu^{21}) normal bioactivity was observed. Particularly noteworthy is the lack of abnormal function when the 3 Glu residues were replaced. An acidic residue in Helix A is thought to be critical for the binding of IL-2, IL-3, and GM-CSF to their respective receptors (23).

(*b*) In predicted Helix C, the $Arg^{103} \rightarrow$ Ala mutein was completely devoid of specific bioactivity in all three bioassay systems (Fig. 3*A*). These results confirm and extend recent work of Grodberg *et al.* (37) who employed only the mouse spleen cell assay (32). Replacement of the adjacent residue Ser^{104} with Ala resulted in normal bioactivity when tested in human UT7 cells, but low (35%) bioactivity when tested in mouse HCD57 cells (Fig. 3*B*). In a similar manner, as shown in Fig. 4*C*, alanine replacement of Leu^{106} in the CD loop resulted in a mutein which had slightly enhanced bioactivity (158%) when tested in the human cells, but low activity when tested in mouse HCD57 cells (48%) or in mouse spleen cells (32%). These results indicate that this region of predicted Helix C contributes strongly to Epo's function. In contrast, Ala replacement of other conserved residues predicted to be on the surface of Helix C (Gln^{92}, Leu^{93}, Lys^{97}, Ser^{100}, Gly^{101}, and Thr^{106}) generated muteins with normal Epo bioactivity. $His^{94} \rightarrow$ Ala and $Asp^{96} \rightarrow$ Ala were not secreted from COS cells as determined by radioimmunoassay or Western blot analyses and therefore could not be assayed for bioactivity.

(*c*) In predicted Helix D, the $Gly^{151} \rightarrow$ Ala mutein was also nearly devoid of specific bioactivity, as assayed in both human UT7 cells and murine HCD57 cells. Moreover, the Ala replacement of the neighboring residue Lys^{152} had decreased specific bioactivity in the two murine cell bioassays but normal activity in human UT7 cells. Of additional interest is the increased specific bioactivity noted in muteins in which nearby residues were replaced by Ala: Asn^{147}, Ser^{146}, Arg^{143}, and perhaps Lys^{140}, Table I). In contrast no significant abnormalities in specific bioactivity was noted when Ala replaced the following residues in Helix D: Asp^{136}, Arg^{139}, and Arg^{150}. Because the exact C-terminal boundary of Helix D is uncertain, and because preliminary data from other laboratories (38, 39) suggest that the C-terminal region of Epo may be functionally important, we

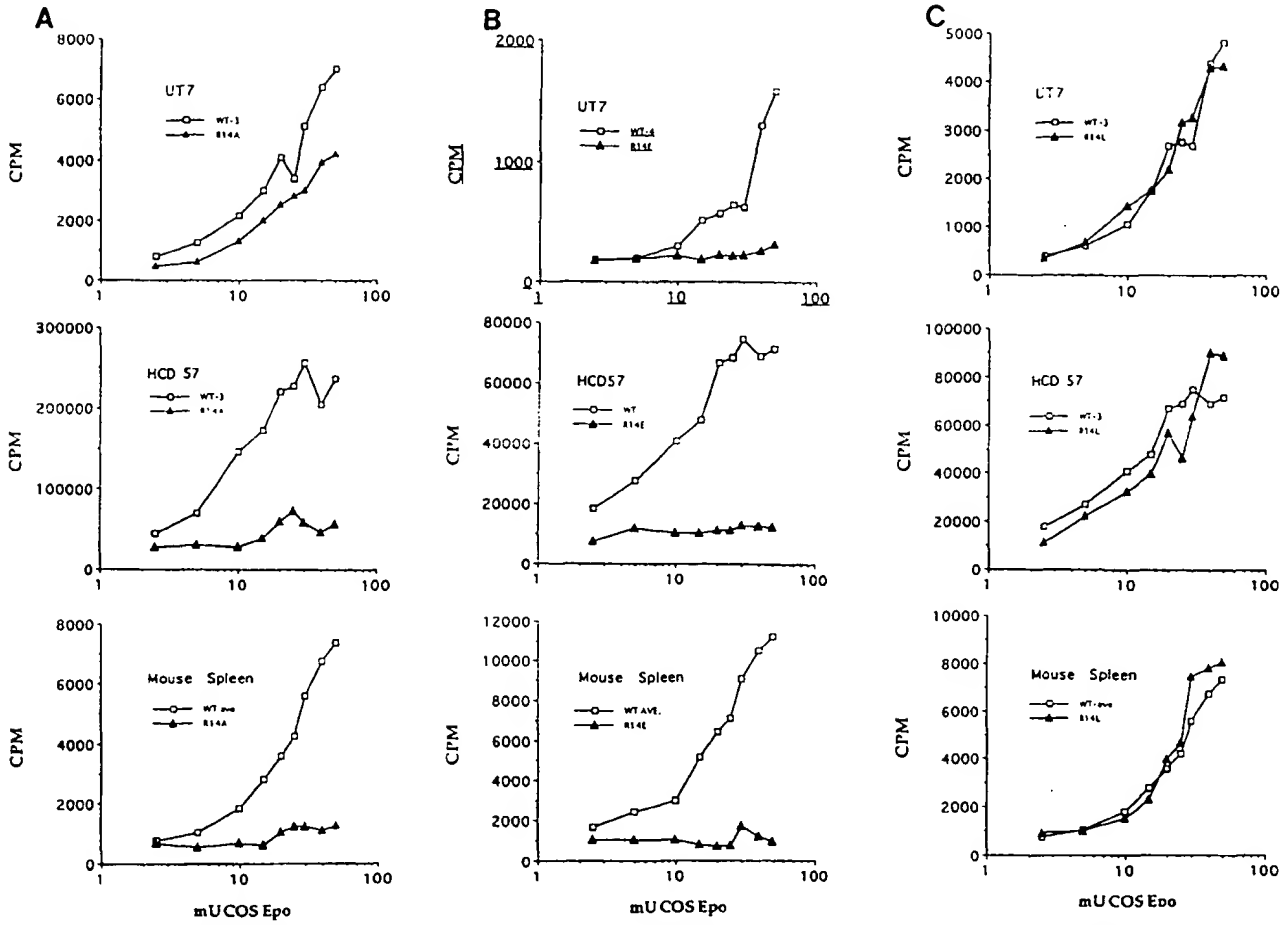


FIG. 2. **Bioassays of mutations at Arg^{14} (*A*, Arg^{14} → Ala; *B*, Arg^{14} → Glu; and *C*, Arg^{14} → Leu).** Epo-dependent HCD57, UT7, and primary murine spleen cells were incubated with increasing amounts of wild type (*WT*) or mutant Epos (muteins). In this, as in subsequent figures, the number after WT refers to a particular transfection. *WT ave* refers to the average of three separate COS cell transfections. The concentrations of wild type Epo and Epo muteins were measured by radioimmunoassay. The uptake of [^{3}H]thymidine provides a measure of Epo-induced cellular proliferation. Differences in incorporation of radioactivity as shown in the *ordinate* reflect different doubling times of the three cell types. The concentration of Epo in the COS cell supernatants is determined by an Epo radioimmunoassay (see "Materials and Methods"). In the bioassay curves the *abscissas* represent increasing concentrations of wild type or mutant Epo. Each assay consisted of three replicate measurements of [^{3}H]thymidine incorporation at nine concentrations of Epo. The area under the curve (*i.e.* the sum of these values, corrected for nonspecific background incorporation) provides the measure of specific bioactivity used in the compilation of Table I.

tested alanine replacements at positions Lys^{153}, Leu^{153}, Lys^{154}, Leu^{155}, Tyr^{156}, Thr^{157}, Gly^{158}, and Glu^{159} which are predicted to be in the adjacent nonhelical segment at the C terminus. As shown in Table I, Lys^{154} → Ala had increased bioactivity, whereas the other replacements were indistinguishable from wild type.

We examined competition between wild type and mutein Epos as shown in Fig. 5. A bioassay in UT7 and in HCD57 cells was performed with a constant amount of wild type Epo and increasing amounts of competitor mutein Epo. Even when present at 5-fold excess, the biologically inactive Arg^{103} → Ala competitor Epo mutein failed to affect the specific biological activity of wild type Epo, in either cell system (Fig. 5*A*). In comparison, a 5-fold excess of the Arg^{14} → Glu mutein resulted in a slight enhancement of bioactivity, suggesting that, at high concentrations, it can bind to the receptor and induce proliferation. As a control (Fig. 5*C*), the Arg^{14} → Leu mutein, which has nearly normal specific bioactivity, when added in 5-fold excess, gave the expected robust increase in proliferation.

At other sites predicted to be on Epo's surface, Ala replacements had no significant effect on biological activity. Lys^{45} → Ala in interhelical region A-B resulted in a mutein with wild type Epo bioactivity. Serine substitutions at Phe^{48}, Ala^{50}, Trp^{51}, and Lys^{52} also resulted in muteins with wild type Epo bioactivity, whereas the Tyr^{49} → Ser mutein had 150% of wild type Epo bioactivity in both UT7 and HCD57 cells. As shown on Table I, Ala replacements of conserved residues predicted to be on the surface of Helix B (Gln^{59}, Glu^{62}, Trp^{64}, Gln^{65}, Gly^{66}, Leu^{69}, Ser^{71}, and Arg^{76}) had normal specific bioactivity. Ala^{73} → Gly had increased bioactivity in UT7 cells. Glu^{72} → Ala was not secreted by COS cells as determined by radioimmunoassay and Western blot analysis (results not shown) and therefore could not be assayed for bioactivity.

DISCUSSION

Initial studies of Epo's structure-function relationships relied primarily on the use of antibodies (40–49). Mapping of specific domains has been based on polyclonal antibodies raised against Epo peptides and monoclonal antibodies raised against either peptides or the intact protein. However, it cannot be assumed that antibodies which neutralize Epo's biological activity always bind to a functionally important domain such as the receptor binding site. Antibodies that bind to a functionally irrelevant site may impair bioactivity by either steric hindrance or the induction of a conformational shift. Reports that antibodies to peptide 111–129 (42, 48) and to the C-terminal

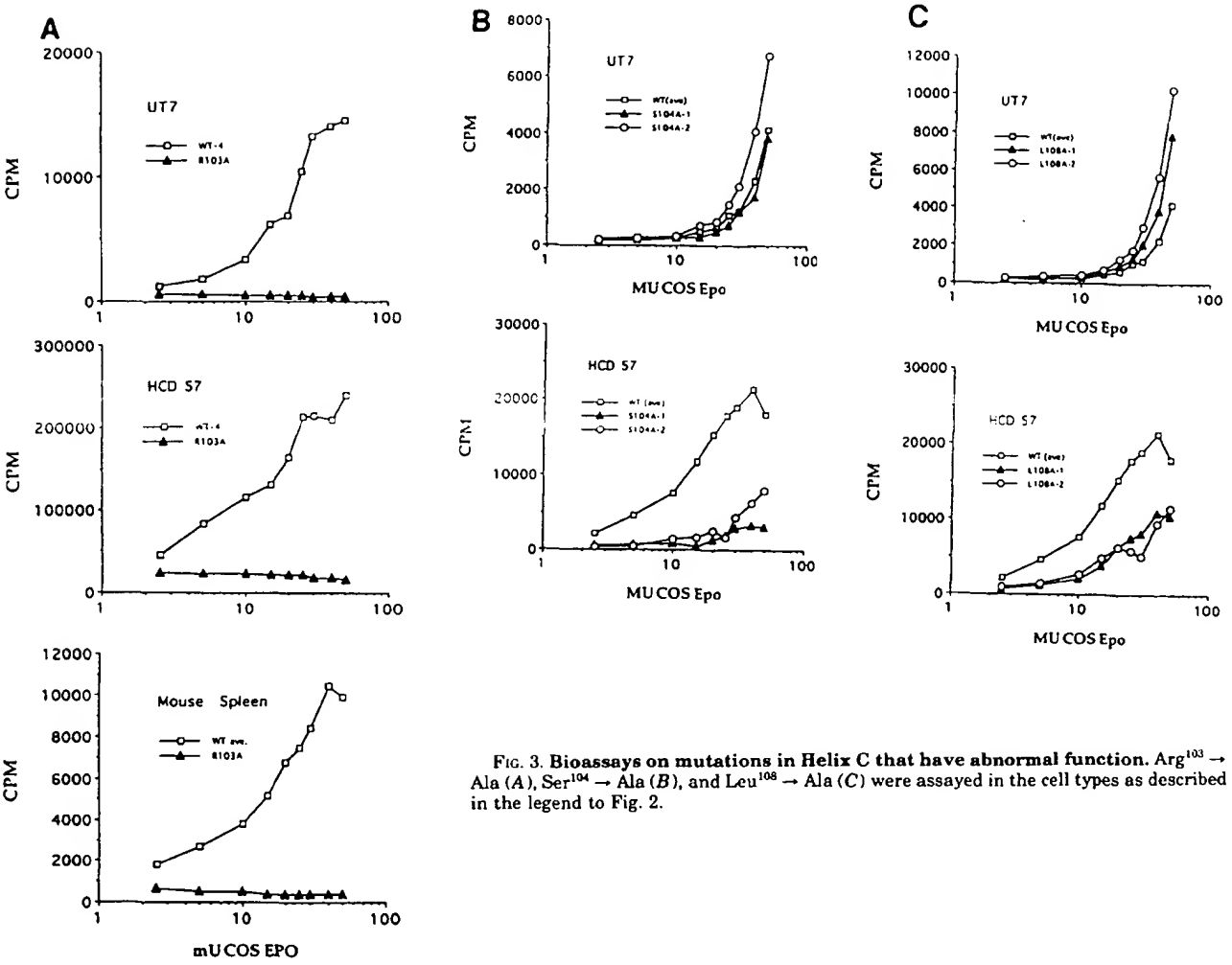


FIG. 3. **Bioassays on mutations in Helix C that have abnormal function.** Arg^{103} → Ala (*A*), Ser^{104} → Ala (*B*), and Leu^{108} → Ala (*C*) were assayed in the cell types as described in the legend to Fig. 2.

peptide 152–166 (47) neutralize Epo function, must be viewed with caution in view of our demonstration that deletions at these sites: Δ110–119, Δ163–166 (25) and Ala replacements at residues 152–159 (Table I) fail to affect biological activity.

In contrast, several (40, 41, 45, 47), but not all (43), investigators have reported that biological activity is not impaired when Epo is bound to antibodies directed against N-terminal peptides. These results must be reconciled with our finding that replacement of Arg^{14} with Ala or Glu, diminishes or abolishes Epo biological activity, respectively. Arg^{14} is predicted to be the fourth residue in α-Helix A, and therefore, the epitope in this portion of the N terminus is likely to differ substantially from the unordered structure of the peptide immunogen. Sytkowski and Fisher (41) reported that a monoclonal antibody raised against the 26-residue N-terminal peptide of human Epo bound to some animal Epos (rat) but not others (sheep and dog). All of these animals share an identical N-terminal sequence through position 15 (24). Thus, this monoclonal antibody is likley to bind an epitope C-terminal to Arg^{14}.

There is much less information in the literature on the preparation and testing of site-directed mutants of Epo. Initially, attention focused on the disulfide bonds (50) (see also Ref. 25) and *N*-linked glycosylation sites (51). Quelle *et al.* (38) prepared muteins in which 6–8-residue cassettes were inserted into the N and C termini of Epo. The former had normal specific bioactivity in agreement with our recent data (25, 52), showing nearly normal specific activity of Epo in which residues 2–5 were deleted and of Epo containing the propeptide at the N terminus. The cassettes which Quelle *et al.* (38) inserted into the C terminus resulted in muteins that were biologically inactive. In contrast, as mentioned above, we have shown that muteins containing deletions of the 4 C-terminal residues (Δ163–166) or replacement of those amino acids by a Lys-Asp-Glu-Lys (KDEL) or polyhistidine sequence, all had full specific bioactivity (25).

Immunochemical studies led Sytkowski and his colleagues (42, 48) to conclude that a functional domain lay within residues 99–129. They then prepared a series of muteins in which Glu-Phe was inserted in place of sequential deletions of 3 adjacent amino acids (53). Muteins in the region 99–110 were not secreted from COS7 cells and, when transcribed and translated *in vitro*, were not biologically active. Recently, this group has reported on Ala replacements of residues 100–109. In full agreement with our results, they found that, with the mouse spleen cell assay, Arg^{103} → Ala was devoid of biological activity, whereas Ser^{104} → Ala and Leu^{108} → Ala had decreased bioactivity. In contrast, our earlier (25) and current results differ almost totally from the recent report of Bittorf *et al.* (39) on four deletion muteins and four Ala replacement muteins. They obtained the following specific bioactivities: Arg^{103} → Ala, 47%; Thr^{106} → Ala, 41%; Lys^{154} → Ala, 9%; and Glu^{159} → Ala, 26%, whereas we obtained values of 5, 114, 160, and 86%, respectively, for these muteins. Equally surprising is their finding that Δ13–17, predicted to be in Helix A, was secreted and had normal biological activity, whereas we found that Δ12–16 was not secreted and devoid of bioactivity (25) and that replace-

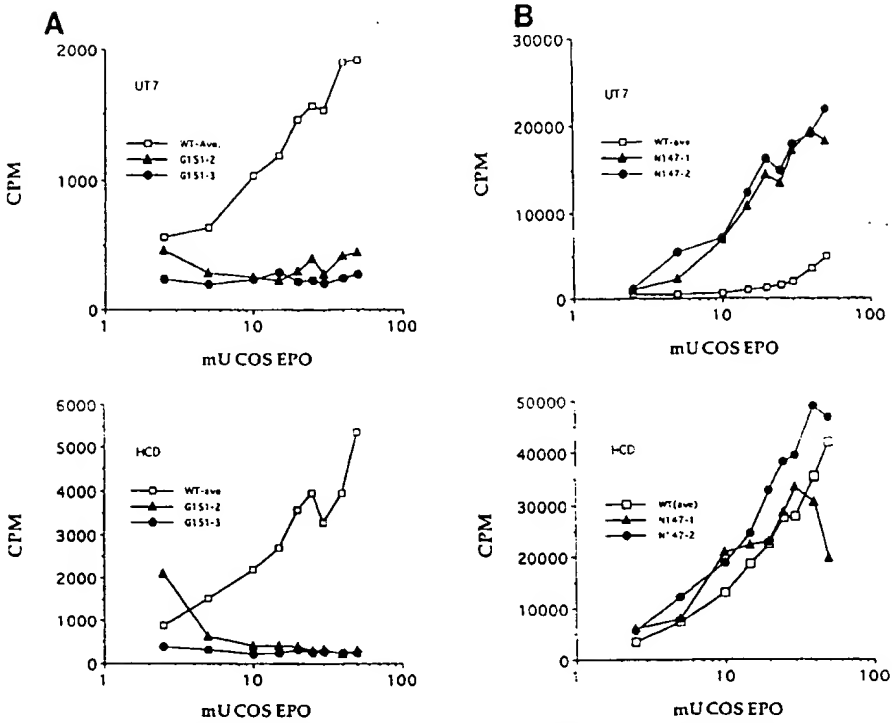


FIG. 4. **Bioassays on mutations in Helix D that have abnormal function.** $Gly^{151} \rightarrow Ala$ (*A*) and $Asn^{147} \rightarrow Ala$ (B) were assayed in the two cell types as described in the legend to Fig. 2.

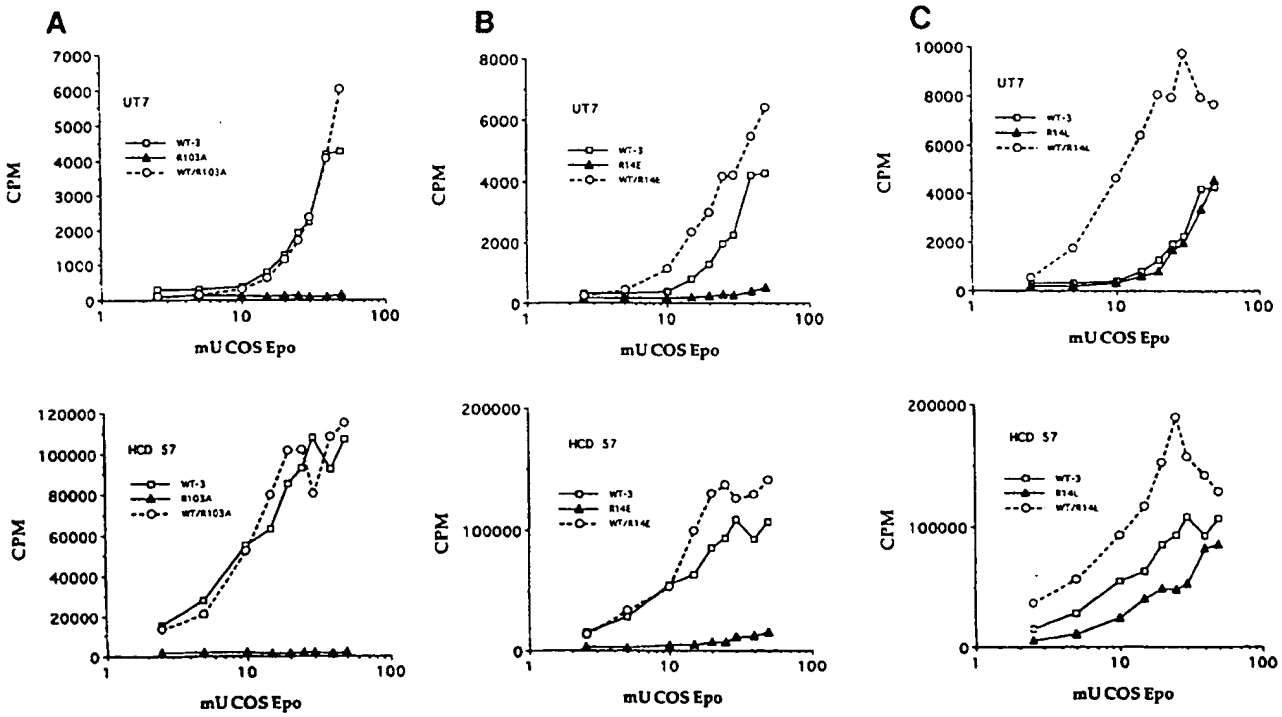


FIG. 5. **Bioassays of mixtures of Epo muteins with wild type Epo.** Each panel shows Epo wild type alone (*open squares*), Epo mutein alone (*solid triangles*), and wild type Epo mixed with a 5-fold excess of mutein (*open circles*). In these mixtures, the *abscissa* refers only to the amount of wild type Epo in the assay. *A*, $Arg^{103} \rightarrow Ala$; *B*, $Arg^{14} \rightarrow Glu$; *C*, $Arg^{14} \rightarrow Leu$.

ments of Arg^{14} markedly impair bioactivity. These investigators (39) employed only the human Epo-sensitive cell line TF-1, which in our hands is unstable and an unreliable assay system for muteins.

Our strategy for designing replacement muteins for studying Epo's structure-function relationships is based a three-dimensional model, reinforced by scanning deletion mutein experiments (25) and mammalian primary sequences (24). As summarized in Table I, we have tested mutations at 54 residues selected by the exclusion criteria described above. The reproducibility, validity, and interpretability of these experiments has been considerably strengthened by the use of three bioas-

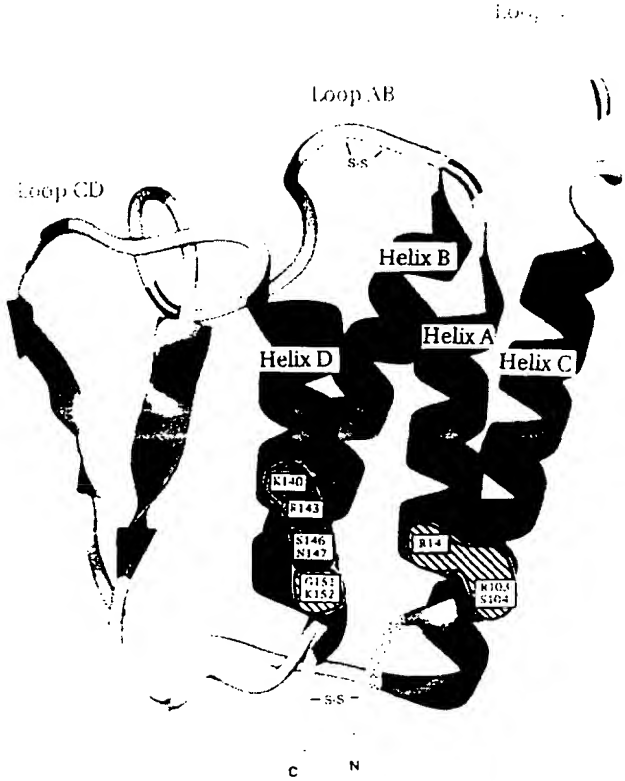


FIG. 6. **Model of predicted three-dimensional structure of Epo showing a proposed functionally important domain that encompasses Arg^{14} on Helix A, Arg^{103} and Ser^{104} on Helix C, and another that encompasses Gly^{151}, Lys^{152}, and perhaps Asn^{147}, Ser^{146}, Arg^{143}, and Lys^{140}.**

say systems, a human cell line, a mouse cell line, and mouse spleen cells.

A number of the muteins that we have tested (R14A, K20A, A73G, S104A, L108A, K140A, R143A, S146A, N147A, K152A) showed higher specific bioactivity with the human cells than with the mouse cells. It is likely that subtle differences in the structures of the human and mouse receptors cooperate with small differences in binding affinity to reveal a functionally abnormal mutein. The competition experiments shown in Fig. 5 suggest that the two muteins with the lowest bioactivity (Arg^{14} → Glu and Arg^{103} → Ala) fail to prevent binding and activity of wild type Epo. This experiment also rules out any nonspecific suppression by these muteins or the COS cell supernatants of the growth of the Epo-dependent assay cells.

Taken together our results show that a restricted number of residues on Helices A, C, and D are involved in Epo's biological activity. We propose the existence of one functional domain (depicted in Fig. 5) encompassing Arg^{14} of Helix A, Arg^{103} and Ser^{104} of Helix C, and perhaps Leu^{108} in the predicted C-D loop. In the three-dimensional model (25) (Fig. 6), Arg^{14} is predicted to be within close proximity of Arg^{103} and Ser^{104}. Moreover, mutagenesis and structural studies on GH, GM-CSF, and IL-4 have demonstrated a functionally important patch, presumably receptor binding, that encompasses portions of Helix A and Helix C (23). We propose a second functional domain (also depicted in Fig. 5) in Helix D at Gly^{151} and perhaps Lys^{152}, where alanine replacements show decreased bioactivity.[3] Our finding of enhanced bioactivities of Ala replacements at nearby residues (Arg^{143}, Ser^{146}, and Asn^{147}) is more difficult to interpret but may indicate an extension of this functional domain. Additional amino acid replacements in the vicinity of these residues in Helices A, C, and D should provide further resolution of our proposed functional domains. We are particularly interested in producing muteins in which a combination of selected replacements result in a maximal enhancement of bioactivity, perhaps enabling the development of an improved therapeutic agent. Further experiments are also in progress to determine if the observed alterations in biological activity of selected muteins can, like GH (54), be explained by parallel changes in the binding affinity to the Epo receptor.

Our structure-function studies suggest that Epo binds to its receptor at the two distinct domains discussed above. Our assignments are consistent with studies on site-directed muteins of other cytokines, including GH, IL-2, IL-3, IL-4, IL-5, and GM-CSF (23), which show that residues on Helix A and Helix D play a critical role in biological activity. X-ray diffraction analysis of the GH·GH receptor complex (13) shows that distinct domains on Helix A and Helix D of GH each bind to a separate receptor molecule, thereby facilitating dimerization of the receptor, perhaps triggering signal transduction.

Acknowledgments—We thank Drs. Fred Cohen, Mark Goldberg, Bob Gruninger, and Linda Mulcahy for valuable discussion. We also thank Drs. David Hankins and Jerry Spivak for providing the HCD 57 cell line and Dr. N. Komatsu for the UT7/Epo cell line.

[3] It is possible that this functional patch also includes a portion of Helix A, analogous to one of the receptor binding domains in GH (13).

REFERENCES

1. Jelkmann, W. (1992) *Physiol. Rev.* **72**, 449–489
2. Porter, D. L., and Goldberg, M. A. (1993) *Exp. Hematol.* **21**, 399–404
3. Miyake, T., Kung, C. K. H., and Goldwasser, E. (1977) *J. Biol. Chem.* **252**, 5558–5564
4. Lai, P.-H., Everett, R., Wang, F.-F., Arakawa, T., and Goldwasser, E. (1986) *J. Biol. Chem.* **261**, 3116–3121
5. Jacobs, K., Shoemaker, C., Rudersdorf, R., Neill, S. D., Kaufman, R. J., Neill, S. D., Kaufman, R. J., Mufson, A., Seehra, J., Jones, S. S., Newick, R., Fritsch, E., Kawakita, M., Shimizu, T., and Miyak, T. (1985) *Nature* **313**, 806–810
6. Lin, F.-K., Suggs, S., Lin, C.-H., Browne, J. K., Smalling, R., Egrie, J., Chen, K. K., Fox, G. M., Martin, F., Stabinsky, Z., Badrawi, S. M., Lai, P.-H., and Goldwasser, E. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 7580–7585
7. Parry, D. A. D., Minasian, E., and Leach, S. J. (1988) *J. Mol. Recognition* **1**, 107–110
8. Bazan, F. (1990) *Immunol. Today* **11**, 350–354
9. Manavalan, P., Swope, D. L., and Withy, R. M. (1992) *J. Protein Chem.* **11**, 321–331
10. Cohen, F. E., and Kuntz, I. D. (1987) *Proteins Struct. Funct. Genet.* **2**, 162–166
11. Curtis, B. M., Presnell, S. R., Srinivasan, S., Sassenfeld, H., Klinke, R., Jeffery, E., Cosman, D., March, C. J., and Cohen, F. E. (1991) *Proteins Struct. Funct. Genet.* **11**, 111–119
12. Abdel-Meguid, S. S., Shieh, H.-S., Smith, W. W., Dayringer, H. E., Violand, B. N., and Bentle, L. A. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 6434–6437
13. de Vos, A. M., Ultsch, M., and Kossiakoff, A. A. (1992) *Science* **255**, 306–312
14. Redfield, C., Smith, L. J., Boyd, J., Lawrence, G. M. P., Edwards, R. G., Smith, R. A. G., and Dobson, C. M. (1991) *Biochemistry* **30**, 11029–11035
15. Powers, R., Garrett, D. S., March, C. J., Frieden, E. A., Gronenborn, A. M., and Clore, G. M. (1992) *Science* **256**, 1673–1677
16. Diederichs, K., Boone, T., and Karplus, A. (1991) *Science* **254**, 1779–1782
17. Walter, M. R., Cook, W. J., Ealick, S. E., Nagabhusan, T. L., Trotta, P. T., and Bugg, C. E. (1992) *J. Mol. Biol.* **224**, 1075–1085
18. Hill, C. P., Osslund, T. D., and Eisenberg, D. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 5167–71
19. Pandit, J., Bohm, A., Jancarick, J., Halenbeck, R., Koths, K., and Kim, S.-H. (1992) *Science* **258**, 1358–1362
20. Bazan, J. F. (1992) *Science* **257**, 410–411
21. McKay, D. B. (1992) *Science* **257**, 412
22. Milburn, M. V., Hassell, A. M., Lambert, M. H., Jordan, S. R., Proudfoot, A. E. I., Graber, P., and Wells, T. N. C. (1993) *Nature* **363**, 172–176
23. Kaushansky, K., and Karplus, P. A. (1993) *Blood* **82**, 3229–3240
24. Wen, D., Boissel, J.-P., Tracy, T. E., Mulcahy, L. S., Czelusniak, J., Goodman, M., and Bunn, H. F. (1993) *Blood* **82**, 1507–1516
25. Boissel, J.-P., Lee, W.-R., Presnell, S. R., Cohen, F. E., and Bunn, H. F. (1993) *J. Biol. Chem.* **268**, 15983–15993
26. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Vol. 1, 2nd Ed., pp. 1.38–1.41, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
27. Kunkel, T. A., Roberts, J. D., and Zakour, R. A. (1987) *Methods Enzymol.* **154**, 367–382
28. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467
29. Kingston, R. E., Chen, C. A., and Okayama, H. (1991) *Current Protocols in Molecular Biology*, Vol. 1, pp. 9.1.1–9.1.9, Green Publishing Associates & Wiley Interscience, NY
30. Goldberg, M. A., Glass, G. A., Cunningham, J. M., and Bunn, H. F. (1987) *Proc. Natl. Acad. Sci. U. S. A.* **84**, 7972–7976

31. Faquin, W. C., Schneider, T. J., and Goldberg, M. A. (1993) *Exp. Hematol.* **21**, 420–426
32. Krystal, G. (1983) *Exp. Hematol.* **11**, 649–60
33. Hankins, W. D., Chin, K., Dons, R., and Szabo, J. (1987) *Blood* **70**, 173a
34. Komatsu, N., Nakauchi, H., Miwa, A., Ishihara, T., Eguchi, M., Moroi, M., Okada, M., Sato, Y., Wada, H., and Yawata, Y. (1991) *Cancer Res* **51**, 341–348
35. Yoshimura, A., D'Andrea, A. D., and Lodish, H. F. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 4139–4143
36. Chou, P. Y., and Fasman, G. D. (1978) *Annu. Rev. Biochem.* **47**, 251–276
37. Grodberg, J., Davis, K. L, and Sytkowski, A. (1993) *Eur. J. Biochem.* **218**, 597–601
38. Quelle, D. E., Lynch, K. J., Burkert-Smith, R. E., Weiss, S., Whitford, W., and Wojchowski, D. M. (1992) *Protein Expression Purif.* **3**, 461–469
39. Bittorf, T., Jaster, R., and Brock, J. (1993) *FEBS Lett.* **336**, 133–136
40. Sue, J. M., and Sytkowski, A. J. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80**, 3651–3655
41. Sytkowski, A. J., and Fisher, J. W. (1985) *J. Biol. Chem.* **260**, 14727–14731
42. Sytkowski, A. J., and Donahue, K. A. (1987) *J. Biol. Chem.* **262**, 1161–1165
43. Wognum, A. W., Lansdorp, P. M., Eaves, C. J., and Krystal, G. (1988) *Blood* **71**, 1731–1737
44. Goto, M., Murakami, A., Akai, K., Kawanishi, G., Ueda, M., Chiba, H., and Sasaki, R. (1989) *Blood* **74**, 1415–1423
45. D'Andrea, A., Szklut, P. J., Lodish, H. F., and Alderman, E. M. (1990) *Blood* **75**, 874–880
46. Wognum, A. W., Lansdorp, P. M., and Krystal, G. (1990) *Exp. Hematol.* **18**, 228–233
47. Fibi, M. R., Stuber, W., Hintz-Obertreis, P., Luben, G., Krumwieh, D., Siebold, B., Zettlmeißl, G., and Kupper, H. A. (1991) *Blood* **77**, 1203–1210
48. Feldman, L., Heinzerling, R., Hillam, R. P., Chern, J. E., Gardiner Frazier, J., Davis, K. L., and Sytkowski, A. J. (1992) *Exp. Hematol.* **20**, 64–68
49. Fibi, M. R., Aslan, M., Hintz-Obertreis, P., Pauly, J. U., Gerken, M., Luben, G., Lauffer, L., Siebold, B., Stuber, W., Nau, G., Zettlmeibl, G., and Harthus, H.-P. (1993) *Blood* **81**, 670–675
50. Lin, F.-K. (1987) *NATO ASI Ser. Ser. H Cell Biol.* **8**, 23–36
51. Dubé, S., Fisher, J. W., and Powell, J. S. (1988) *J. Biol. Chem.* **263**, 17516–17521
52. Boissel, J.-P., Wen, D., and Bunn, H. F. (1994) *Ann. N. Y. Acad. Sci.* **718**, 203–212
53. Chern, Y., Chung, T., and Sytkowski, A. J. (1991) *Eur. J. Biochem.* **202**, 225–229
54. Wells, J. A. (1994) *Curr. Opin. Cell Biol.* **6**, 163–173

Mapping of the Active Site of Recombinant Human Erythropoietin

By Steve Elliott, Tony Lorenzini, David Chang, Jack Barzilay, and Evelyne Delorme

Recombinant human erythropoietin (rHuEPO) variants have been constructed to identify amino acid residues important for biological activity. Immunoassays were used to determine the effect of each mutation on rHuEPO folding. With this strategy, we could distinguish between mutations that affected bioactivity directly and those that affected bioactivity because the mutation altered rHuEPO conformation. Four regions were found to be important for bioactivity: amino acids 11 to 15, 44 to 51, 100 to 108, and 147 to 151. EPO variants could be divided into two groups according to the differential effects on EPO receptor binding activity and in vitro biologic activity. This suggests that rHuEPO has two separate receptor binding sites. Mutations in basic residues reduced the biologic activity, whereas mutations in acidic residues did not. This suggests that electrostatic interactions between rHuEPO and the human EPO receptor may involve positive charges on rHuEPO.

ERYTHROPOIETIN (EPO) is a hormone produced primarily by the kidney. It is involved in the growth and maturation of erythroid cells from precursors. Decreased production of EPO due to kidney failure results in anemia. The human EPO gene has been cloned,[1,2] and the hormone has been produced by recombinant DNA techniques (recombinant human EPO [rHuEPO]). The human gene encodes a 165–amino acid mature protein. The secreted protein contains approximately 40% carbohydrate and has an approximate molecular weight of 30 kD.[3] The carbohydrate has been shown to have an effect on protein stability and solubility, but it is not required for in vitro bioactivity.[4,5] This indicates that the active site is retained in the protein portion of the molecule.

The mechanism by which EPO stimulates erythropoiesis is partly understood. EPO binds to specific cell-surface EPO receptors. Activation of the EPO receptor triggers intracellular signaling events including phosphorylation of the receptor, followed by activation of the JAK-STAT, RAS, and PI3 kinase pathways.[6-11] These signaling pathways trigger cells to undergo proliferation and differentiation and to block apoptosis. Cloning of the murine and human EPO receptors[12,13] and later identification of a constitutively active receptor mutant with an Arg129 to Cys mutation[14,15] led to the proposal that dimerization of the receptor is responsible for its activation. This proposal was supported by the observation that bivalent monoclonal antibodies but not monovalent Fab fragments could activate the EPO receptor in the absence of EPO.[16]

If activation of the EPO receptor is indeed due to homodimerization, then EPO must be capable of simultaneously binding two receptors. The fact that rHuEPO exists in solution as a monomer suggests that it must have two receptor binding sites much like that of growth hormone. In this case, a growth hormone monomer activates its receptor by cross-linking two growth hormone receptors.[17-19] However, for EPO, the location of the two binding sites has not been definitively identified.

Several strategies have been used to try to identify the active site of EPO. These studies included mapping of neutralizing antibody epitopes[20-26] and analysis of the effect of deletions[27,28] or point mutations[21,29-32] on bioactivity. These studies implicate several regions that are important for bioactivity in EPO, Arg 14, residues between 100 and 110, and some residues near the C-terminus. Interpretation of the results of these experiments is difficult, because the crystal or nuclear magnetic resonance structures of EPO have not been reported. In addition, interpreting mutagenesis results is complicated by the difficulty of distinguishing between changes in bioactivity caused by changes in the active site and changes in protein folding that indirectly affect bioactivity. The effect of mutations on folding of EPO variants was not established in previous studies, except for an Arg103 to Ala mutation.[31]

We wished to accurately map the active site of rHuEPO and determine whether one or two binding sites could be identified. We constructed 141 different rHuEPO variants and tested them for biologic activity. We have previously reported the development of an immunologic approach that allowed testing of the effect of mutations on rHuEPO conformation.[20,21] Analysis of the effect of mutations in rHuEPO on its immunoreactivity with this method has been used previously to identify buried residues in rHuEPO and those important for the rHuEPO structure. With this strategy, we have identified mutations that are either predicted to be on the exposed surface of rHuEPO or only modestly affect the rHuEPO structure. Thus, we have identified residues that appear to be part of the rHuEPO active site. According to models that suggest that rHuEPO is a four-helical bundle,[20,27,33,34] the residues important for bioactivity can be localized to two different regions. This suggests that rHuEPO has two different binding sites.

MATERIALS AND METHODS

Construction and expression of rHuEPO variants. The rHuEPO cDNA was obtained from COS-1 cells transformed with a human rHuEPO genomic clone.[2,35] The coding region was recovered as an 810-nucleotide *Bst*EII-*Bgl* II DNA fragment and was cloned into an SV40 vector.[4] The vector contained a dihydrofolate reductase (DHFR) gene for selection in Chinese hamster ovary (CHO) DHFR- cells. All the variants were expressed in COS-1 cells. Transfection was by either the calcium phosphate method[20,36] or electroporation.[20] Conditioned medium was collected after 3 to 5 days. Aliquots were made and stored at −70°C. In cases for which larger amounts of material were needed, stable CHO cell transformants were made.

From Amgen, Thousand Oaks, CA.
Submitted February 24, 1996; accepted September 3, 1996.
Address reprint requests to Steve Elliott, PhD, Amgen, Amgen Center, Thousand Oaks, CA 91320.
The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. section 1734 solely to indicate this fact.

0006-4971/97/8902-0018$3.00/0

CHO cells were transfected by the calcium phosphate method, and transformants were selected for DHFR expression.

To construct rHuEPO variants, in vitro oligo-directed mutagenesis was performed.[20] In brief, a DNA segment containing human EPO coding sequences was transferred to M13mp18, and single-stranded DNA was produced in a *dut ung Escherichia coli* strain (RZ1032). The single-stranded DNA was annealed with synthetic oligonucleotides containing the desired mutations, and was then extended with Klenow polymerase in the presence of T4 DNA ligase. The DNA was then transfected into *E. coli* strain JM109. Variants with the desired mutations were identified by differential hybridization using the mutant oligonucleotide. The presence of the mutations was confirmed by DNA sequencing, and then the DNA segment was transferred to the expression vector.

Analysis of rHuEPO expression. The rHuEPO concentration (units per milliliter) in conditioned medium for each rHuEPO variant was determined by four to six different immunoassays. The EPO standard used in the assays was rHuEPO produced in CHO cells. ^{125}I-EPO was obtained from Amersham (Arlington Heights, IL). The isolation and characterization of antibodies and immunoassays has been described elsewhere.[20,21] These immunoassays included a radioimmunoassay (RIA) using an antibody raised to purified rHuEPO (RIA-P); a RIA using an antibody raised to a synthetic peptide corresponding to the first 20 amino acids of rHuEPO (RIA-N); and enzyme-linked immunosorbent assays (ELISAs) using two different monoclonal antibodies that recognize different rHuEPO epitopes (EIA-D11 and EIA-F12). The ELISA format involved capture of rHuEPO by the monoclonal antibody, and then an incubation with one of two different signal polyclonal antibodies raised to rHuEPO. Two different polyclonal signal antibodies raised in goats or rabbits were available. The RIA-N immunoassay is largely unaffected by rHuEPO conformation. In addition, the epitopes recognized by these immunoassays have been determined.[20] Therefore, it was possible to determine the rHuEPO concentration in samples by selecting assays that were unaffected by the introduced mutation. Each rHuEPO variant was also assayed with the monoclonal antibody 9G8A (RIA-9G8A). The specific 9G8A immunoreactivity was determined as units of 9G8A immunoreactive EPO divided by total units of immunoreactive EPO as determined earlier.

Increased immunoreactivity measured by RIA-9G8A is indicative of altered folding. Completely denatured rHuEPO yields an immunoreactivity measured by this antibody that was 20- to 40-fold higher than expected.[21] In addition, immunoreactivity appears to increase in proportion to the degree of unfolding.[20] Immunoreactivity measured by this assay that was 2.5 times higher than expected was considered indicative of altered folding.

Purification of rHuEPO variants. Conditioned medium from CHO cells expressing the various rHuEPO variants was purified as previously described.[37]

In vitro bioassays. In vitro bioactivity was determined by measuring thymidine uptake in a factor-dependent cell line, 32D,[38,39] that had been made dependent on EPO for growth by introduction of a murine EPO receptor gene (32D + EPOR[40]). Assays were performed as previously described[40] with the following modifications. 32D cells were factor-starved for 3 to 5 hours in assay medium (RPMI 1640 medium; GIBCO BRL, Grand Island, NY) containing 10% fetal bovine serum and 1% Glutamine Pen-Strep (Irvine Scientific, Santa Ana, CA) lacking EPO. Test samples or EPO standard (rHuEPO) were added to wells in a 96-well microtiter plate; 50 μL starved cells were then added (15,000 cells/well), and the plates were incubated in a humidified incubator at 37°C and 6% CO_2. After 18 to 24 hours, 50 μL methyl-^{3}H-thymidine (1 mCi/mL, 20 Ci/mmol) diluted 1:100 in assay medium was added. Plates were incubated for an additional 4 hours at 37°C and 6% CO_2. Labeled cells were harvested onto glass fiber filtermats using a PHD cell harvester (Cambridge Technology Inc, Watertown, MA) and deionized water as a washing solution. Filters were rinsed a final time with 2-propanol, and then were dried and counted in a Beckman (Fullerton, CA) model LS6000IC scintillation counter. Biologic activity was expressed in units using rHuEPO as the standard. The specific bioactivity was units of bioactive EPO divided by units of immunoreactive EPO.

Receptor binding. Receptor binding activity was determined by cold-displacement assays with high-specific activity ^{125}I-EPO (Amersham) on a human cell line (OCIM1[41]). Cells were grown to 2 to 5 × 10^5/mL, collected by centrifugation, washed two times in binding buffer (RPMI 1640 medium/1% bovine serum albumin/25 mmol/L HEPES, pH 7.3), and then resuspended at 1 to 2 × 10^7/mL in binding buffer containing 0.1% azide and 10 μg/mL cytochalisin B. Cells were incubated with varying amounts of sample and ^{125}I-EPO at 37°C. After 3 hours, cells were centrifuged through phthalate oil (60:40 vol/vol dibutyl:dinonyl phthalate). The tubes were then quick-frozen in dry ice ethanol, the pellets were clipped, and the amount of bound ^{125}I-EPO was determined by counting in a gamma counter. Receptor binding activity was determined by comparison of the displacement by the rHuEPO variants to a cold-displacement standard curve generated with a rHuEPO standard. Values were thus relative to the standard and are reported in units per milliliter. Specific receptor binding activity was determined by dividing by the amount of immunoreactive EPO determined by immunoassay.

RESULTS

Construction and analysis of rHuEPO variants. We have previously constructed over 200 different rHuEPO mutations. These have been analyzed to determine which ones affect rHuEPO structure.[20] We have also used immunologic data to predict residues thought to be on the solvent-exposed surface of rHuEPO. Thus, we selected from this group 141 individual mutations with substitutions at 118 of 165 amino acid positions in rHuEPO and tested them for in vitro biological activity. At several positions, multiple substitutions were made and the rHuEPO variants were assayed. Thus, the additional mutations would confirm the importance of the position for bioactivity and provide information on the amino acid requirements at that site.

Each variant was transiently expressed in COS-1 cells, and conditioned medium was collected. In several cases, stable expression was obtained in CHO cells to produce larger quantities of rHuEPO variants. To perform the analysis on a large scale, most of the variants were not purified. Instead, each supernatant was assayed by up to six different immunoassays. Therefore, it was possible to determine the concentration of variants in the cell-conditioned medium using the assays that were unaffected by the introduced mutation. The effect of each mutation on rHuEPO conformation was also estimated with RIA-9G8A. The immunoreactivity measured by this assay increases when rHuEPO is denatured.[21]

Figure 1 shows results of the in vitro bioassays that measured thymidine uptake in the EPO-dependent cell line 32D + EPOR. Figure 1 also shows the effect of each mutation on folding according to the results of RIA-9G8A immunoassays. Several amino acid segments in rHuEPO could be mutated without affecting bioactivity. These included residues 21 to 44, 52 to 95, and 109 to 140. Thus, these regions do not appear to play a major role in bioactivity. In addition, a termination codon introduced at Thr163 also had no apparent

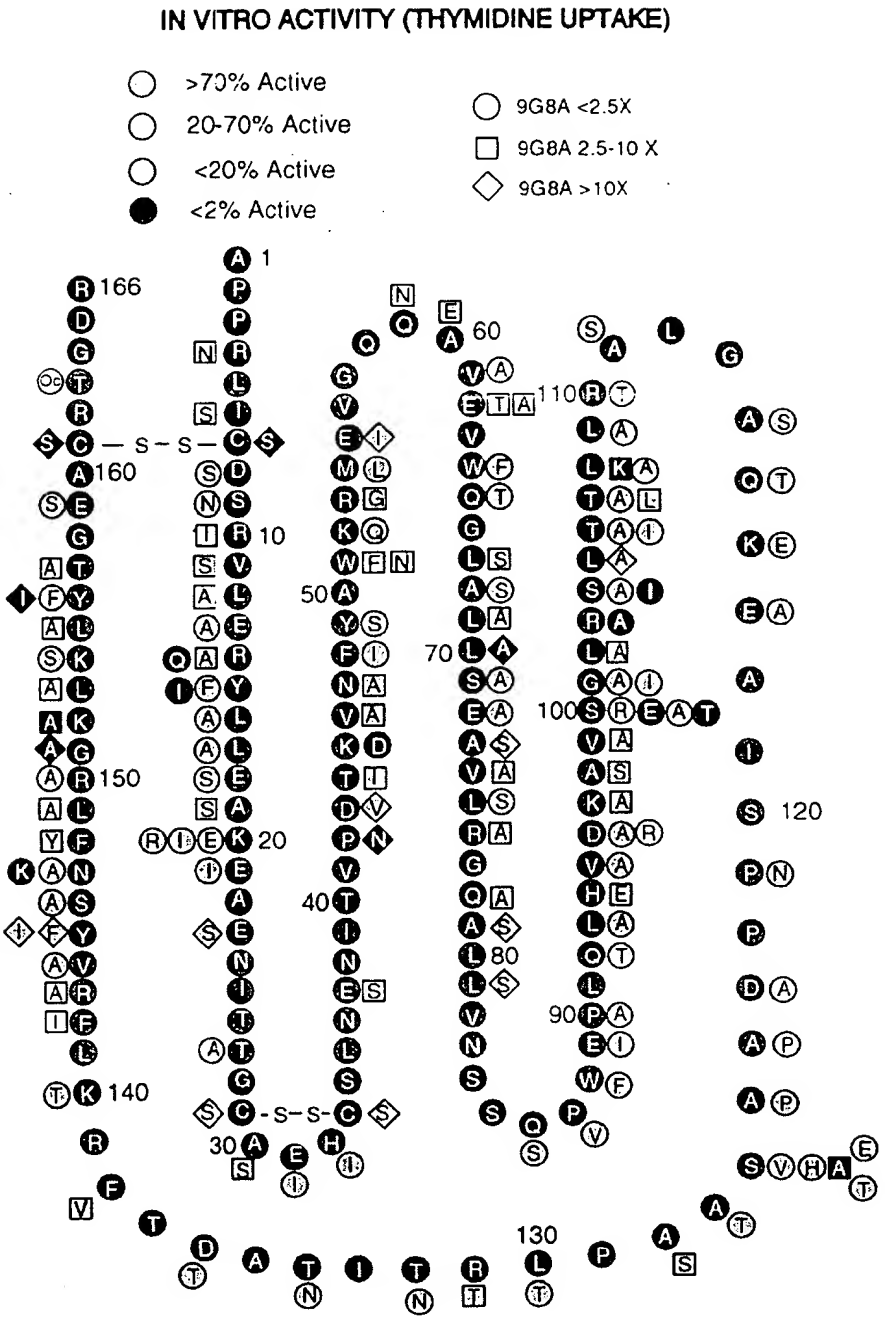


Fig 1. Mutagenesis of rHuEPO. Amino acid residues in the rHuEPO sequence (black circles with white letters) and rHuEPO variants (adjacent to rHuEPO sequences with shaded symbols) are indicated. O°, substitution of a termination codon, TGA. The shape of the symbols indicates the effect of the mutation on RIA-9G8A immunoreactivity. The shading code represents the effect on in vitro bioactivity.

effect on bioactivity, which suggests that the C-terminal residues are not important for bioactivity. Introduction of a proline at −2 relative to the amino-terminal Ala affects cleavage of the signal peptide, resulting in a three–amino acid extension (Val Pro Gly[20,42,43]). In vitro activity and folding measured by RIA-9G8A with a Pro-2 variant was normal (data not shown), which suggests that the amino terminus is not important for bioactivity. This variant has an altered amino terminus as evidenced by a lack of immunoreactivity with antibodies that bind to the amino terminus.[20] Forty-three mutations at 36 positions did have an affect on bioactivity. These mutations were concentrated in four regions: residues 10 to 20, 44 to 51, 96 to 108, and 142 to 156. Mutations that reduced the bioactivity by over 50-fold were found in all four regions. Mutations at several other scattered positions also reduced bioactivity. These included the mutations

Gln59, Glu62, Leu67, and Leu70 and disulfide bonds. These latter mutations also affect folding and/or are predicted to be in buried positions.[20] Thus, the mutations in these residues may affect bioactivity indirectly.

All rHuEPO variants in Fig 1 were also tested for the ability to stimulate formation of erythroid bursts using erythroid precursors from human bone marrow. The results were qualitatively similar to those in Fig 1 (data not shown). Mutations in all the residues that reduced thymidine uptake in 32D + EPOR also reduced erythroid burst formation. In addition, variants identified as having the greatest reduction in thymidine uptake activity also had the greatest reduction in erythroid burst formation. The fact that human cells were used in this assay suggests that the residues identified as being important for activation of 32D + EPOR (murine cells) are also important for activation of human cells. Thus, the mutations in the putative active site affect both proliferation and differentiation and are important in both species.

Role of Tyr and acidic and basic residues in bioactivity. It has been reported that chemical modification of tyrosines can inactivate rHuEPO.[44,45] To see if we could identify which tyrosines were responsible for this effect, each of the four Tyr residues were individually mutated. Mutations at Tyr49 or Tyr145 had no effect on bioactivity (Fig 1). The Tyr15 and Tyr156 mutations included Phe and Ile substitutions. The Phe15 and Phe156 mutations had no effect on either bioactivity or rHuEPO folding, as evidenced by a normal immunoreactivity measured by RIA-9G8A. However, an Ile156 substitution affected both RIA-9G8A immunoreactivity (20-fold increase) and bioactivity (11-fold decrease), and an Ile15 substitution reduced bioactivity over 200-fold while at the same time having no apparent effect on RIA-9G8A. These results suggest that an aromatic amino acid is required at position 15 for bioactivity, and that an aromatic amino acid at position 156 may be required for proper folding.

Chemical modification of amino groups (eg, Lys) has also been reported to affect bioactivity.[44,45] In addition, electrostatic interactions between charged residues in four-helix-bundle proteins and their receptors have been reported to play a role in hormone receptor recognition.[46] We therefore examined the effect of substitutions in the charged residues in rHuEPO. rHuEPO has six Asp, 12 Glu, eight Lys, 13 Arg, and two His residues.[1,2] All the charged residues except Arg162 were individually mutated to uncharged and/or oppositely charged residues. The Asp at 165 was removed by substituting a termination codon at Thr163 (Fig 1). Some substitutions of the acidic residues affected folding. However, only a Glu62 to Thr mutation resulted in a decrease in bioactivity twofold. This variant also had a 7.5-fold increase in RIA-9G8A immunoreactivity, and an Ala62 mutation had no effect on bioactivity. This suggests that Glu62 may be important for structure but not for bioactivity, and that acidic residues as a group do not have a major role to play in bioactivity. The two His residues at 32 and 94 were also mutated. Mutations in these residues also had no effect on bioactivity. The results were different when the lysine and arginine residues were mutated. Mutations at four of eight lysine residues (positions 20, 45, 97, and 152) and five of 13 arginine residues (positions 10, 14, 103, 143, and 150) affected bioactivity. Some of the mutations that affected bioactivity also affected folding (eg, at Lys97, Lys152, Arg10, and Arg143). Therefore, we cannot eliminate the possibility that for these latter four residues the reduced bioactivity was due to indirect effects. Mutations at Arg14, Arg103, Arg150,

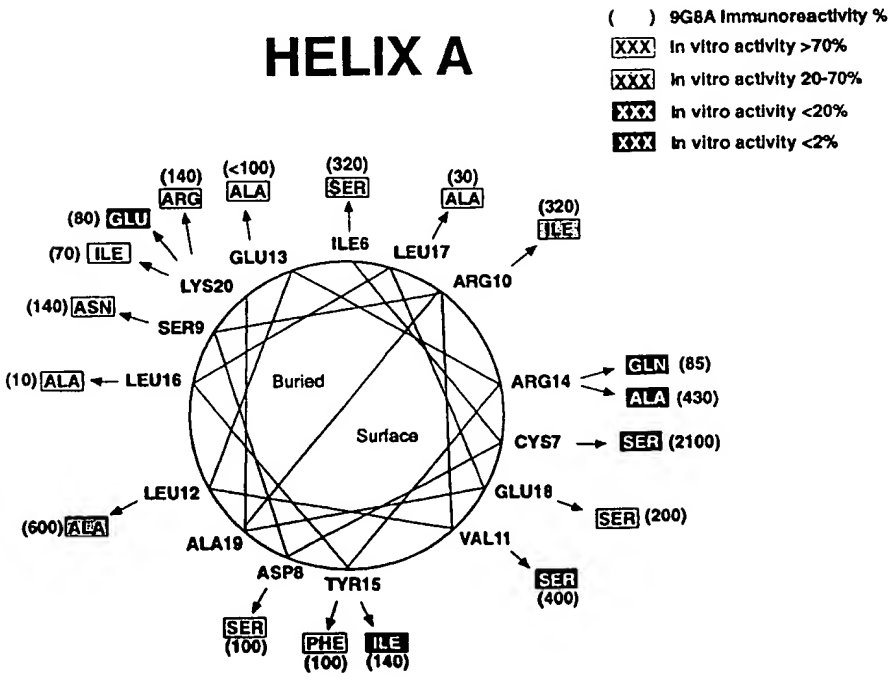


Fig 2. Effect of mutations on structure and function for residues 5 to 23. A helical wheel presentation shows amino acids found in rHuEPO (closest to the circle). Arrows point to mutations (in boxes). RIA-9G8A immunoreactivity for each variant is shown in parentheses adjacent to the mutation. Shading indicates the effect of the mutation on in vitro bioactivity.

and Lys45 had dramatic effects on bioactivity with no apparent effect on folding. This suggests that these are part of the active site. Substitutions at Lys20 revealed additional information about charge requirements. An isoleucine substitution at Lys20 had no effect on bioactivity; however, a glutamic acid substitution reduced bioactivity approximately 11-fold. Receptor binding activity was also reduced eightfold for the Glu20 mutation (data not shown). None of these mutations significantly affected RIA-9G8A immunoreactivity. This suggests that there is not a requirement for a positively charged residue at this position, but that a negatively charged residue disrupts receptor binding and thus bioactivity. This observation in combination with the fact that only mutations in the positively charged lysine and arginine residues affected bioactivity suggests that the electrostatic interactions between rHuEPO and its receptor may be modulated primarily by positive charges on rHuEPO and presumably by negative charges on the EPO receptor.

Identification of active site mutants. With an analysis of this type, one would expect many of the mutations to result in a reduction in biologic activity because the rHuEPO structure was affected. For example, disruption of disulfide bonds has been shown to alter rHuEPO structure and activity, and thus they presumably are not part of the active site.[21,27,32] Some of these latter types of mutations were identified according to high immunoreactivity as measured by RIA-9G8A. For example, at position 126, five different substitutions were evaluated. Only an Ala126 mutation affected RIA-9G8A immunoreactivity, and this was the only mutation that affected biologic activity even though the other substitutions were nonconservative. Thus, it is unlikely that Ser126 is part of the active site.

Some of the regions in which mutations affect biologic activity are in potential α-helices.[20] Mutations that disrupt protein structure as evidenced by increased RIA-9G8A immunoreactivity are concentrated on the buried sides of the predicted helices. The predicted helices have been designated helix A (Fig 2), helix B (Fig 3), helix C (Fig 4), and helix D (Fig 5). Thus, it is possible to discriminate between direct and indirect effects of the mutations on activity by examining the locations of the affected residues on helical wheels.

Mutations in three positions in helix A (Fig 2) had large effects on bioactivity and were concentrated on the same side of the helix: Val11, Arg14, and Tyr15. Gln14 and Ile15 mutations did not affect immunoreactivity measured by RIA-9G8A but reduced the bioactivity over 50-fold, which suggests that they may be folded normally. Therefore, these residues are probably part of the active site. The Ser11 mutation also affected bioactivity, and it is on the same face of the predicted helix as Arg14 and Tyr15. However, the mutation had a fourfold increased immunoreactivity measured by RIA-9G8A, suggesting an effect on folding. Mutations in residues Arg6, Arg10, and Leu12 had modest effects on bioactivity. Mutations in these residues had an effect on rHuEPO folding as evidenced by their increased RIA-9G8A immunoreactivity, and they are on the same face of the helix as the bulk of the hydrophobic amino acids (eg, Leu12, Leu16, and Leu17). Therefore, these mutations may be in residues that are buried. The Cys7 to Ser mutation had an effect on biologic activity, but it also had a dramatic effect on RIA-9G8A immunoreactivity, presumably because it is involved in formation of a Cys7-Cys161 disulfide bond that is important for rHuEPO structure.[20,27]

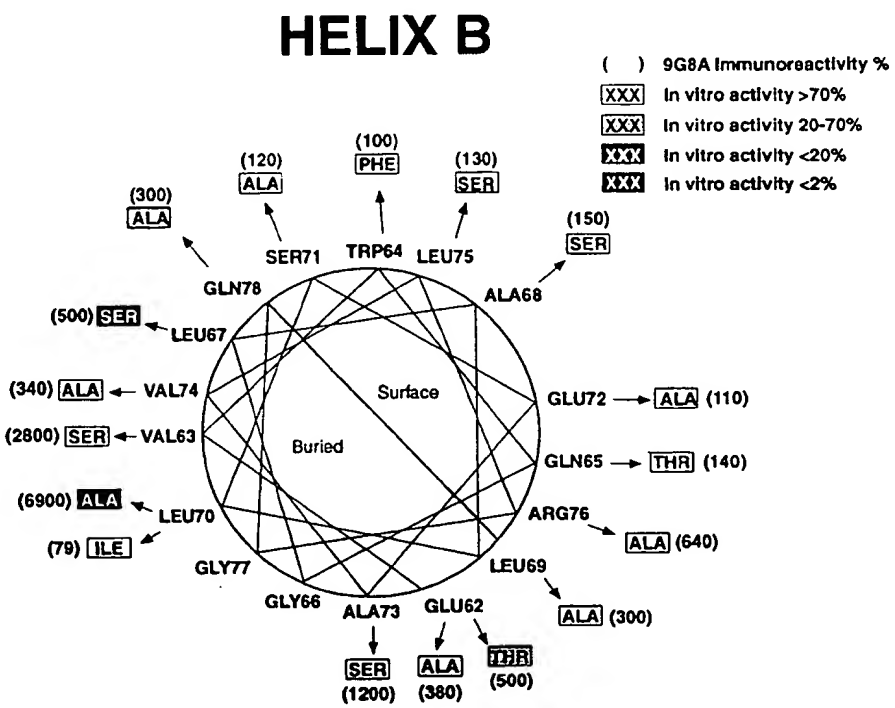


Fig 3. Effect of mutations on structure and function for residues 62 to 79. A helical wheel presentation shows amino acids found in rHuEPO (closest to the circle). Arrows point to mutations (in boxes). RIA-9G8A immunoreactivity for each variant is shown in parentheses adjacent to the mutation. Shading indicates the effect of the mutation on in vitro bioactivity. Predicted buried and exposed surfaces are indicated.

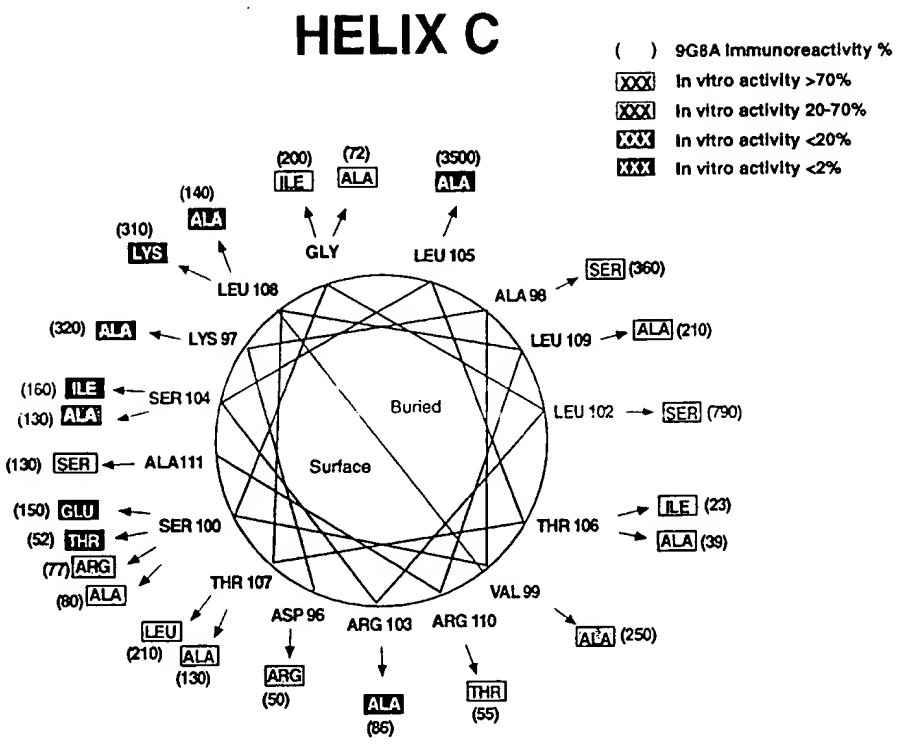


Fig 4. Effect of mutations on structure and function for residues 94 to 111. A helical wheel presentation shows amino acids found in rHuEPO (closest to the circle). Arrows point to mutations (in boxes). RIA-9G8A immunoreactivity for each variant is shown in parentheses adjacent to the mutation. Shading indicates the effect of the mutation on in vitro bioactivity. Predicted buried and exposed surfaces are indicated.

Mutations at three positions in helix B affected bioactivity: Glu62, Leu67, and Leu70. However, mutations in all three positions affected RIA-9G8A immunoreactivity and are predicted to be in buried positions. This suggests that mutations in helix B that affect bioactivity do so indirectly (Fig 3).

Ser100, Arg103, Ser104, and Leu108 (helix C, Fig 4) are predicted to be on the protein surface, and mutations in them appear to be folded normally but dramatically affect bioactivity. Four substitutions were examined at Ser100, including Glu, Thr, Arg, and Ala. Only the Ala100 mutation had normal bioactivity. Mutations in residues on the opposite face of the predicted helix, at Leu102 and Leu105, severely affect folding as evidenced by high RIA-9G8A. Thus, it is unlikely that Leu102 or Leu105 are part of the active site. On the predicted exposed surface of helix D (Fig 5), mutations at Asn147, Arg150, and Gly151 had the greatest effect on bioactivity. An Arg143 to Ala mutation had a modest effect on bioactivity, but this mutation also affected folding. A Leu155 to Ala mutation also affected bioactivity. This residue may be partially buried or may be important for structure, since mutations in it had an effect on folding. An Ala151 substitution resulted in greater than a 150-fold decrease in bioactivity. However, the Ala151 mutation affected folding significantly. Several mutations in predicted buried positions affected RIA-9G8A immunoreactivity and bioactivity: Phe142, Leu149, Lys152, Leu153, and Tyr156.

Mutations between residues 42 and 51 also affected bioactivity. This region is thought to be in a connecting loop between helices A and B. An Asp45 mutation resulted in reduced bioactivity but appeared to be folded normally, suggesting that it may play a role in bioactivity (Table 1). The substitutions Asn42, Ile44, Ala46, Asn51, and Phe51 affected bioactivity also. However, the mutations also affected folding, so we cannot definitively conclude which ones are part of the active site.

Receptor binding activity and biologic activity of rHuEPO variants reveal two classes. The residues thought to play a role in in vitro bioactivity were examined to see what effect they had on binding to the EPO receptor. Several rHuEPO variants were expressed in both COS-1 and CHO cells (Table 1). All the mutants that affected bioactivity in the A/B connecting loop are shown, even though we do not know which of them may be directly involved in biologic activity. The helix A, C, and D variants shown include only those with mutations in residues thought to be on the solvent-exposed surface. We purified three different rHuEPO variants and natural-sequence rHuEPO to confirm the methods used to calculate activities. This was necessary because we determined the rHuEPO concentration in conditioned medium by immunoassay. The results are shown in Table 2. In all cases, results on purified material compared favorably with results obtained with unpurified material. This was true for effects on both biologic activity and receptor binding activity.

Some of the mutations had modest effects on receptor binding activity but no detectable biologic activity (Tables 1 and 2). Similar results were observed for both COS and CHO cell-expressed rHuEPO variants. This result suggests that rHuEPO binding by itself is not sufficient for activation of the EPO receptor. The four regions important for bioactiv-

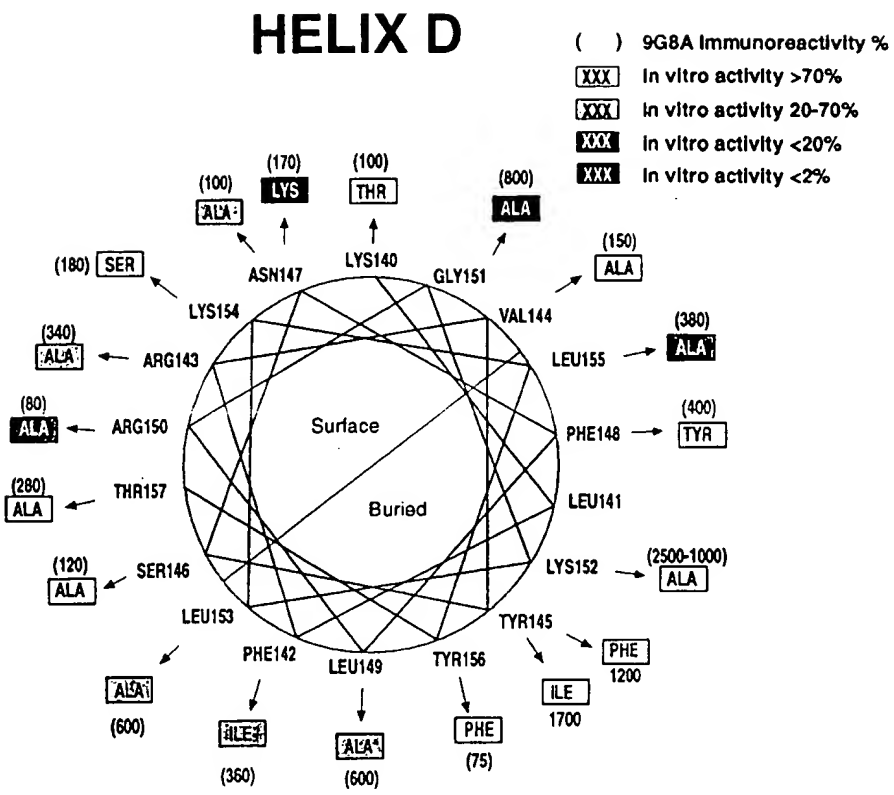


Fig 5. Effect of mutations on structure and function for residues 140 to 157. A helical wheel presentation shows amino acids found in rHuEPO (closest to the circle). Arrows point to mutations (in boxes). RIA-9G8A immunoreactivity for each variant is shown in parentheses adjacent to the mutation. Shading indicates the effect of the mutation on in vitro bioactivity. Predicted buried and exposed surfaces are indicated.

ity could be divided into two groups according to the differential effects on biologic activity and receptor binding activity. The variants with changes in helices A and C retained substantial receptor binding activity (≥5% of normal) even though bioactivity in many cases was low or not detectable (Arg14/Gln, Tyr15/Ile, Ser100/Glu, Ser100/Thr, Arg103/Ala, Ser104/Ile, and Lys108/Ala; Table 1). Thus, a relatively small change in receptor binding activity due to mutations in these regions resulted in much greater effects on bioactivity. In contrast, substitutions in residues between 42 and 51 and in residues in helix D had receptor binding activity that more closely correlated with biologic activity (Val46/Ala, Trp51/Phe, Lys147/Ala, and Arg150/Ala). Mutations in these two regions also differed from mutations in helices A and C in that some mutations resulted in dramatic decreases in receptor binding activity (≥100-fold decrease; Lys45/Asp, Asn147/Lys, and Gly151/Ala). One explanation for these observations is that the two groups may represent residues in different receptor binding sites, and each receptor binding site has a different role to play in activity.

DISCUSSION

We have devised a strategy to evaluate the effect on folding and bioactivity of 141 different rHuEPO mutations at 118 positions, and found that four regions were important for bioactivity. These have been localized to helix A (Val11, Arg14, and Tyr15), helix C (Ser100, Arg103, Ser104, and Leu108), helix D (Asn147, Arg150, Gly151, and Leu155), and the A/B connecting loop (residues 42 to 51).

Previous studies implicated three of four regions we identified as having an important role to play in biological activity. These regions included residues between 10 and 15, 100 and 108, and 150 and 159.[30-32] In particular, alanine-scanning mutagenesis data suggested that amino acids 14, 103, 104, 105, 108, and 151 were important for bioactivity and that residues 100, 101, 102, 106, 107, and 109 were not.[31,32] We have also introduced alanines in all these positions and obtained qualitatively similar results. At three of these positions, other substitutions were made. These substitutions, Glu100/Ser, Ser104/Ile, and Leu108/Lys mutations, reduced bioactivity over 1,000-, 500-, and 5,000-fold, respectively, whereas Ala substitutions resulted in modest or no effects. Thus, at these positions Ala can substitute for the existing amino acid, resulting in effective activation of the EPO receptor, but other amino acids cannot. Our results and those of Wen et al[32] and Grodberg[31] are in disagreement with another study that reported that an Ala103 mutation resulted in a modest effect on bioactivity.[30] An Ala154 mutation was reported to reduce bioactivity 10-fold.[30] We examined Arg154 and Ser154 substitutions and found no effect on bioactivity. In addition, Wen et al[32] tested an Ala154 mutation and reported no effect on bioactivity. The same group[30] reported that a Glu159 to Ala mutation reduced bioactivity fourfold. We tested a Ser159 mutation and found no effect on bioactivity. The reasons for the discrepancies are not known, but may be due to the different amino acid substitutions, the different methods of determining EPO concentration, or the methods of expression and/or storage.

A recent analysis of hematopoietic cytokines and their receptors suggests that electrostatic interactions are important for receptor binding. Complementary electrostatic interactions between the binding surfaces of growth hormone and IL-4 and their corresponding receptors have been suggested.[46] In addition, mutations in both positively charged and negatively charged residues in growth hormone affect activity.[18] We examined the effect of mutations in acidic and basic residues, and found that only mutations in the basic residues affected bioactivity of rHuEPO. This differs from what is suggested for growth hormone, which uses both positive and negative charges for interaction with the growth hormone receptor. However, the difference in the importance of negatively charged residues for rHuEPO and human growth hormone is consistent with the difference in the charge of growth hormone and rHuEPO; growth hormone has a net charge of −4e, and the protein portion of rHuEPO has a net charge of +5e. EPO is more like IL-4, which has a net charge of +7e. In this case, the charge in the IL-4 receptor at the predicted IL-4 receptor-hormone interface is primarily negative.[46] This suggests that the positively charged (basic) residues on rHuEPO may interact with the negatively charged (acidic) residues on the EPO receptor.

We and others have reported models of rHuEPO structure based on the four–helix-bundle motif.[20,27,33] The four regions important for bioactivity identified by ourselves and others[30-32] map to two sites on the rHuEPO models. These include site 1, containing the residues between amino acids 41 and 52 and amino acids 147, 150, and 151. The second site includes residues 11, 14, 15, 100, 103, 104, and 108. Mutations in site 1 have substantial effects on both receptor binding and in vitro biological activity. Mutations in site 2 have modest effects on receptor binding activity and much greater effects on in vitro biological activity.

This pattern of active site residues is analogous to that of growth hormone, which has two binding sites, exists as a monomer in solution, and also has a four–helix-bundle structural motif.[19]

Mutations in the two different sites in rHuEPO and human growth hormone behave similarly with respect to receptor binding activity and biologic activity. The observations can be explained by a sequential binding model, whereby the residues in site 1 have high affinity and are responsible for the initial interaction with the human growth hormone receptor. The residues in site 2 have less effect on receptor binding, but are required for dimerization. One consequence of the human growth hormone site 2 mutations is that they can act as antagonists.[47] We have found that mutations in site 2 in rHuEPO can also act as antagonists (data not shown). Also consistent with this proposal is the recent report that the extracellular domain of the EPO receptor and rHuEPO can form a 2:1 complex in solution.[48] In addition, both high- and low-affinity binding can be detected. Thus, it appears that EPO receptor dimerization may proceed in a manner similar to that of growth hormone.

Table 1. Activity of rHuEPO Variants

Location	Mutation	Expression* (U/mL)	Receptor Binding† (% of total)	In Vitro Activity‡ (% of total)	RIA-9G8A§ (% of total)
	Human EPO	67	130	187	100
Helix A	Val11/Ser	45	25	6.3	380
	Arg14/Ala	44	40	11	590
	Arg14/Gln	21	28	1.7	82
	Tyr15/Ile	11	28	<0.43	150
A/B Loop	Pro42/Asn	0.8	<1	<0.01	1,600
	Thr44/Ile	7	52	8.7	640
	Lys45/Asp	53	0.2	<0.03	180
	Val46/Ala	9	20	7.8	774
	Typ51/Phe	97	40	34	340
Helix C	Ser100/Glu	79	8	<0.01	150
	Ser100/Thr	14	11	<0.2	58
	Arg103/Ala	7.7	20	<0.12	86
	Ser104/Ile	30	7.3	<0.2	100-240
	Ser104/Ala	12	36	12	150
	Leu108/Lys	35	10	<0.02	330
Helix D	Asn147/Lys	5	1.2	0.3	260
	Arg150/Ala	47	19	8	97
	Gly151/Ala	29	<0.7	<0.03	760
	Leu155/Ala	42	48	16	439

* The rHuEPO constructs were transfected into Cos cells. After 3 to 5 days, the concentration (U/mL) of rHuEPO was immunoassay (see text).

† Receptor binding assays were performed using OCIM-1 cells (human erythroleukemia cells) and cell conditioned media from COS cells transfected with the wild-type or mutant EPO constructs. The assay measures displacement of ^{125}I rHuEPO. Specific activity was calculated by dividing by the concentration determined by immunoassay. Results are expressed as a percentage of total.

‡ In vitro activity was the ability of cell conditioned media containing wild-type or mutant EPO to stimulate ^{3}H thymidine uptake in 32D+EPOR (32D cells transfected with a murine EPO receptor). Specific activity was calculated by dividing by the concentration determined by immunoassay. Results are expressed as a percentage of total.

§ 9G8A assays were radioimmunoassays using the anti-EPO monoclonal antibody 9G8A. Specific activity was calculated by dividing by concentration determined by immunoassay. Results are expressed as a percentage of total. Increased immunoreactivity is indicative of altered folding.

Table 2. Activity of Purified rHuEPO Variants

Mutation	Receptor Binding U/mg	Receptor Binding % of rHuEPO	In Vitro Activity U/mg	In Vitro Activity % of rHuEPO	RIA-9G8A (U/mg)
rHuEPO	175,000	100	179,000	100	144,000
Arg14/Gln	32,000	18	6,100	3.4	193,000
Trp51/Phe	40,200	23	54,000	30	454,000
Asn147/Lys	1,600	0.92	980	0.55	500,000

Activity was determined as described in Table 1 and the Materials and Methods. Specific activity was calculated from activity (U/mL) by dividing by the protein concentration (determined from absorbance at A280).

ACKNOWLEDGMENT

We wish to thank T. Papayannopoulou (University of Washington, Seattle, WA) for OCIM1 cells; R. Pacifici for 32D + EPOR cells; B. Aguero, J. Grant, D. Greene, and J. Egrie for expert technical assistance and helpful discussions; K. Chen, S. Suggs, A. Harcourte,

J. Katzowitz, A. Janssen, and L. Antonio for DNA sequence analysis; K. Aoki and S. Asher for purification of rHuEPO variants; and T. Jones and B. Goodman for synthetic oligonucleotides.

REFERENCES

1. Jacobs K, Shoemaker C, Ruderdorf R, Neill SD, Kaufman RM, Mufson A, Seehra J, Jones SS, Hewick R, Fritch EF, Kawakita M, Shimizu T, Miyake T: Isolation and characterization of genomic and cDNA clones of human erythropoietin. Nature 313:806, 1985

2. Lin FK, Suggs S, Lin CH, Browne JK, Smalling R, Egrie JC, Chen KK, Fox GM, Martin F, Stabinsky Z, Badrawi S, Lai PH, Goldwasser E: Cloning and expression of the human erythropoietin gene. Proc Natl Acad Sci USA 82:7580, 1985

3. Davis JM, Arakawa T, Strickland TW, Yphantis DA: Characterization of recombinant human erythropoietin produced in Chinese hamster ovary cells. Biochemistry 26:2633, 1987

4. Delorme E, Lorenzini T, Giffin J, Martin F, Jacobsen F, Boone T, Elliott S: Role of glycosylation on the secretion and biological activity of erythropoietin. Biochemistry 31:9871, 1992

5. Narhi LO, Arakawa T, Aoki KH, Elmore R, Rohde MF, Boone T, Strickland TW: The effect of carbohydrate on the structure and stability of erythropoietin. J Biol Chem 266:23022, 1991

6. Miura O, D'Andrea AD, Kabat D, Ihle JN: Induction of tyrosine phosphorylation by the erythropoietin receptor correlates with mitogenesis. Mol Cell Biol 11:4895, 1991

7. Ihle JN, Witthuhn BA, Quelle FW, Yamamoto K, Thierfelder WE, Kreider B, Silvennoinen O: Signaling by cytokine receptor superfamily: JAKs and STATs. Trends Biochem Sci 19:222, 1994

8. Witthuhn BA, Quelle FW, Silvennoinen O, Yi T, Tang B, Miura O, Ihle JN: JAK2 associates with the erythropoietin receptor and is tyrosine phosphorylated and activated following stimulation with erythropoietin. Cell 74:227, 1993

9. Torti M, Marti KB, Altschuler D, Yamamoto K, Lapetina EG: Erythropoietin induces $p21^{ras}$ activation and p120GAP tyrosine phosphorylation in human erythroleukemia cells. J Biol Chem 267:8293, 1992

10. Damen JE, Mui AL-F, Puil L, Pawson T, Krystal G: Phosphatidylinositol 3-kinase associates, via its SH2 domains, with the activated erythropoietin receptor. Blood 81:3204, 1993

11. Mayeux P, Dusanter-Fourt I, Muller O, Mauduit P, Sabbah M, Druker B, Vainchenker W, Fisher S, Lacombe C, Gisselbrecht S: Erythropoietin induces the association of phosphatidylinositol 3'-kinase with a tyrosine-phosphorylated protein complex containing the erythropoietin receptor. Eur J Biochem 216:821, 1993

12. D'Andrea AD, Lodish HF, Wong GG: Expression cloning of the murine EPO receptor. Cell 57:277, 1989

13. Jones S, D'Andrea AD, Haines LL, Wong GG: Human erythropoietin receptor: Cloning, expression, and biologic characterization. Blood 76:31, 1990

14. Watowich SS, Yoshimura A, Longmore GD, Hilton DJ, Yoshimura Y, Lodish HF: Homodimerization and constitutive activation of the erythropoietin receptor. Proc Natl Acad Sci USA 89:2140, 1992

15. Yoshimura A, Longmore G, Lodish HF: Point mutation in the exoplasmic domain of the erythropoietin receptor resulting in hormone-independent activation and tumorigenicity. Nature 348: 647, 1990

16. Elliott S, Lorenzini T, Yanagihara D, Chang D, Elliott G: Activation of the erythropoietin (EPO) receptor by bivalent anti-EPO receptor antibodies. J Biol Chem 271:24691, 1996

17. Cunningham BC, Ultsch M, DeVos AM, Mulkerrin MG, Clauser KR, Wells JA: Dimerization of the extracellular domain of the human growth hormone receptor by a single hormone molecule. Science 254:821, 1991

18. Cunningham BC, Wells JA: High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. Science 244:1081, 1989

19. De Vos AM, Ultsch M, Kossiakoff AA: Human growth hormone and extracellular domain of its receptor: Crystal structure of the complex. Science 255:306, 1992

20. Elliott S, Lorenzini T, Chang D, Delorme E, Giffin J, Hesterberg L: Fine-structure epitope mapping of antierythropoietin monoclonal antibodies reveals a model of recombinant human erythropoietin protein structure. Blood 87:2702, 1996

21. Elliott S, Lorenzini T, Boone T, Chang D, Delorme E, Dunn C, Egrie J, Giffin J, Talbot C, Hesterberg L: Isolation and characterization of conformation sensitive anti-erythropoietin monoclonal antibodies: Effect of disulfide bonds and carbohydrate on rHuEPO structure. Blood 87:2714, 1996

22. Fibi MR, Stuber W, Hintz-Obertreis P, Luben G, Krumwieh D, Siebold B, Zettlmeißl G, Kupper HA: Evidence for the location of the receptor-binding site of human erythropoietin at the carboxyl-terminal domain. Blood 77:1203, 1991

23. Egrie J, Lane J: Use of monoclonal and polyclonal antipeptide antibodies to assay and characterize erythropoietin. Fed Proc 43:1892, 1984

24. Feldman L, Heinzerling R, Hillam RP, Chern YE, Frazier JG, Davis KL, Sytowski AJ: Four unique monoclonal antibodies to the putative receptor binding domain of erythropoietin inhibit the biological function of the hormone. Exp Hematol 20:64, 1992

25. Sytowski AJ, Fisher JW: Isolation and characterization of an anti-peptide monoclonal antibody to human erythropoietin. J Biol Chem 260:14727, 1985

26. Sytowski AJ, Donahue KA: Immunochemical studies of human erythropoietin using site-specific anti-peptide antibodies. J Biol Chem 262:1161, 1987

27. Boissel JP, Lee WR, Presnell SR, Cohen FE, Bunn HF: Erythropoietin structure-function relationships. J Biol Chem 268:15983, 1993

28. Chern Y, Chung T, Sytowski AJ: Structural role of amino acids 99-110 in recombinant human erythropoietin. Eur J Biochem 202:225, 1991

29. Bill RM, Winter PC, McHale CM, Hodges VM, Elder GE, Caley J, Flitsch SL, Bicknell R, Lappin TRJ: Expression and mutagenesis of recombinant human and murine erythropoietins in *Escherichia coli*. Biochim Biophys Acta 1261:35, 1995

30. Bittorf T, Jaster R, Brock J: Structural and functional characteristics of recombinant human erythropoietin analogues. FEBS Lett 336:133, 1993

31. Grodberg J, Davis KL, Sytowski AJ: Alanine scanning mutagenesis of human erythropoietin identifies four amino acids which are critical for biological activity. Eur J Biochem 218:597, 1993

32. Wen D, Boissel J-P, Showers M, Ruch BC, Bunn HF: Erythropoietin structure-function relationships: Identification of functionally important domains. J Biol Chem 269:22839, 1994

33. Haniu M, Narhi LO, Arakawa T, Elliott S, Rohde MF: Recombinant human erythropoietin (rHuEPO): Cross-linking with disuccinimidyl esters and identification of the interfacing domains in EPO. Protein Sci 2:1441, 1993

34. Bazan JF: Haemopoietic receptors and helical cytokines. Immunol Today 11:350, 1990

35. Law ML, Cai G-Y, Lin F-K, Wei Q, Huang S-Z, Hartz JH, Morse H, Lin C-W, Jones C, Kao F-T: Chromosomal assignment of the human erythropoietin gene and its DNA polymorphism. Proc Natl Acad Sci USA 83:6920, 1986

36. Wigler M, Pellicer A, Silverstein S, Axel R: Biochemical transfer of single-copy eucaryotic genes using total cellular DNA as donor. Cell 14:725, 1978

37. Elliott S, Bartley T, Delorme E, Derby P, Hunt R, Lorenzini T, Parker V, Rohde MF, Stoney K: Structural requirements for addi-

tion of O-linked carbohydrate to recombinant erythropoietin. Biochemistry 33:11237, 1996

38. Greenberger JS, Sakakeeny MA, Humphries RK, Eaves CJ, Eckner RJ: Demonstration of permanent factor-dependent multipotential (erythroid/neutrophil/basophil) hematopoietic progenitor cell lines. Proc Natl Acad Sci USA 80:2931, 1983

39. Ihle JN, Peppersack L, Rebar L: Regulation of T cell differentiation: In vitro induction of 20 alpha-hydroxysteroid dehydrogenase in splenic lymphocytes from athymic mice by a unique lymphokine. J Immunol 126:2184, 1981

40. Pacifici RE, Thomason AR: Hybrid tyrosine kinase/cytokine receptors transmit mitogenic signals in response to ligand. J Biol Chem 269:1571, 1994

41. Broudy VC, Lin N, Egrie J, De-Haen C, Weiss T, Papayannopoulou T, Adamson JW: Identification of the receptor for erythropoietin on human and murine erythroleukemia cells and modulation by phorbol ester and dimethyl sulfoxide. Proc Natl Acad Sci USA 85:6513, 1988

42. Lin F-K, Lin C-H, Lai P-L, Browne JK, Egrie JC, Smalling R, Fox GM, Chen KK, Castro M, Suggs S: Monkey erythropoietin gene: Cloning, expression and comparison with the human erythropoietin gene. Gene 44:201, 1986

43. Lin FK: The molecular biology of erythropoietin, in Rich IN (ed): Molecular and Cellular Aspects of Erythropoietin, vol H8. Berlin, Germany, Springer-Verlag, 1987, p 23

44. Goldwasser E: Erythropoietin and red cell differentiation, in Cunningham D, Goldwasser E, Watson J (eds): Control of Cellular Division and Development, part A. New York, NY, Liss, 1981, p 487

45. Satake R, Kozutsumi H, Takeuchi M, Asano K: Chemical modification of erythropoietin: An increase in in vitro activity by guanidination. Biochim Biophys Acta 1038:125, 1990

46. Demchuk E, Mueller T, Oschkinat H, Sebald W, Wade RC: Receptor binding properties of four-helix-bundle growth factors deduced from electrostatic analysis. Protein Sci 3:920, 1994

47. Fuh G, Cunningham BC, Fukunaga R, Nagata S, Goeddel DV, Wells JA: Rational design of potent antagonists to the human growth hormone receptor. Science 256:1677, 1992

48. Philo JS, Aoki KH, Arakawa T, Narhi LO, Wen J: Dimerization of the extracellular domain of the erythropoietin (EPO) receptor by EPO: One high-affinity and one low-affinity interaction. Biochemistry 35:1681, 1996

Efficiency of signalling through cytokine receptors depends critically on receptor orientation

Rashid S. Syed*, Scott W. Reidt, Culwei Li*, Janet C. Cheetham*, Kenneth H. Aoki*, Beishan Liut, Hangjun Zhant, Timothy D. Osslund*, Arthur J. Chirino*, Jiandong Zhang*, Janet Finer-Mooret, Steven Elliott*, Karen Sitney*, Bradley A. Katzt, David J. Matthewst, John J. Wendoloski*, Joan Egrie* & Robert M. Stroudt‡

* Amgen Inc., One Amgen Center Drive, Thousand Oaks, California 91320-1789, USA

† Axys Pharmaceuticals Inc., 180 Kimball Way, South San Francisco, California 94080, USA

‡ Department of Biochemistry and Biophysics and Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, California 94143-0448, USA

Human erythropoietin is a haematopoietic cytokine required for the differentiation and proliferation of precursor cells into red blood cells[1]. It activates cells by binding and orientating two cell-surface erythropoietin receptors (EPORs) which trigger an intracellular phosphorylation cascade[2]. The half-maximal response in a cellular proliferation assay is evoked at an erythropoietin concentration of 10 pM (ref. 3), 10^{-2} of its K_d value for erythropoietin–EPOR binding site 1 ($K_d \approx 1$ nM), and 10^{-5} of the K_d for erythropoietin–EPOR binding site 2 ($K_d \approx 1\ \mu M$)[4]. Overall half-maximal binding (IC_{50}) of cell-surface receptors is produced with ~0.18 nM erythropoietin, indicating that only ~6% of the receptors would be bound in the presence of 10 pM erythropoietin. Other effective erythropoietin-mimetic ligands that dimerize receptors can evoke the same cellular responses[5,6] but much less efficiently, requiring concentrations close to their K_d values (~0.1 μM). The crystal structure of erythropoietin complexed to the extracellular ligand-binding domains of the erythropoietin receptor, determined at 1.9 Å from two crystal forms, shows that erythropoietin imposes a unique 120° angular relationship and orientation that is responsible for optimal signalling through intracellular kinase pathways.

A soluble analogue of *Escherichia coli*-expressed erythropoietin (Swiss Prot accession no. P01588) was made by substituting lysine residues at three *N*-linked glycosylation sites (N24K, N38K and N83K). The extracellular ligand-binding domain of the erythropoietin receptor (EPObp; Swiss-Prot accession number P19235) was expressed in *Pichia pastoris* cells, with the *N*-linked glycosylation site removed by an N52Q mutation. An N164Q mutation was included to remove the potential for deamidation at the normal -N-G- sequence[7]. An A211E mutation was introduced for better expression and potentially enhanced folding efficiency on the basis of its role in improving the transportation of EPOR from the endoplasmic reticulum to the cell surface[8].

Erythropoietin–$(EPObp)_2$ complex crystals diffracting to 2.8 Å ($a = 73.3$ Å, $b = 80.3$ Å, $c = 134.9$ Å, $P2_12_12_1$; Form 1) were obtained and the structure was determined by single isomorphous replacement (SIR) based on a mercury derivative, followed by solvent density modification by the methods of Abrahams[9] in the SOLOMON algorithm (Table 1). This was alternated with density averaging of equivalent pairs of individual receptor domains. Higher-resolution data (1.9 Å) were obtained from a new crystal form, which used an erythropoietin analogue with two additional mutations, P121N and P122S, introduced into the CD loop. These changes were suggested on the basis of ^{15}N NMR relaxation data, which showed extremely low-order parameters for the residues (Glu 117–Ala 128) in this loop with conformational heterogeneity in the backbone in the vicinity of the proline residues[10]. This analogue of erythropoietin, when complexed with EPObp, crystallized in a new unit cell ($a = 58.4$ Å, $b = 79.3$ Å, $c = 136.5$ Å, $P2_12_12_1$; Form 2) with a diffraction limit beyond 1.9 Å. The crystal structure of the erythropoietin–$(EPObp)_2$ complex (Form 2) was solved by multiple isomorphous replacement with anomalous scattering (MIRAS) using one uranium and two platinum derivatives (Table 1).

The structure of human erythropoietin complexed with EPObp

Table 1 Statistics for data collection and structure refinement.

Parameters	Native (Form 1)	$HgAc_2$ (Form 1)	Native (Form 2)	$UO_2(NO_3)_2$ (Form 2)	K_2PtCl_4 (Form 2)	$K_2Pt(NO_2)_4$ (Form 2)
Data collection						
Resolution (Å)	2.8	2.8	1.9	2.3	2.6	2.5
Total observations	88,794	105,221	251,681	252,332	167,497	137,452
Unique reflections	20,677	19,853	49,480	28,936	22,589	21,321
Average I/σ	8.6	9.3	19.1 (2.5)	12.0 (3.9)	12.7 (5.0)	12.4 (3.7)
Completeness (%)	91.6 (89.2)	88.0 (85.3)	97.0 (82.9)	99.7 (97.5)	99.9 (99.8)	93.4 (83.8)
R_{sym} (%)*	12.2 (37.4)	9.8 (38.2)	4.4 (37.7)	7.8 (46.9)	9.2 (46.9)	7.7 (44.8)
Phase determination						
Soaking conc. (mM)		1 and 10		0.1	0.1	4
Soaking time (h)		16		10	11	24
Number of sites		2		2	2	2
R_{iso} (%)†		7.3		7.4	12.7	12.0
R_c (%)‡		64.3		65.0/43.0	62.0/44.0	65.0/44.0
Phasing power§		1.5/1.0		1.5/1.6	1.7/1.5	1.3/1.4
Refinement						
Resolution (Å)	80.0–2.8		50.0–1.9			
Reflections, $F > 2\sigma$	18,797		44,507			
R_{cryst}‖/R_{free}¶ (%)	19.0/33.0		24.2/31.8			
No. water mols	0		300			
R.m.s.d. bond length (Å)	0.008		0.006			
R.m.s.d. bond angle (deg)	1.39		1.3			
B-factor (Å²)	37.0		34.3			

Values given in parentheses are for highest-resolution shell. R.m.s.d., root-mean-square deviation.

* $R_{sym} = \Sigma|I_{obs} - I_{mean}|/\Sigma I_{obs}$, unweighted *R*-factor on *I* among symmetry mates.

† $R_{iso} = \Sigma||F_{ph}| - |F_p||/\Sigma|F_{ph}|$, where F_p and F_{ph} are the structure factor amplitudes of the native and derivative data respectively.

‡ $R_c = \Sigma||F_{ph}| \pm |F_p|| - |F_H|/\Sigma|F_{ph}| \pm |F_p||$ for centric reflections. The second entry is for anomalous dataset where $R_c = \Sigma||F^+_{ph}| - |F^+_{obs}|| + ||F^-_{ph}| - |F^-_{obs}||/\Sigma|F^+_{ph}| + |F^-_{ph}|$ calculated for the top 25% largest Bijvoet differences. Data cutoff >2.3 Å and $F/\sigma = 4.0$.

§ Phasing power = r.m.s. $(|F_H|/E)$, where F_H is the calculated heavy-atom structure factor amplitude and E is the residual lack of closure error. The second entry is for anomalous phasing power, r.m.s. $(|F_H|/E_{anom})$, where F_H is the anomalous correction component amplitude. Data cutoff >2.3 Å and $F/\sigma = 4.0$.

‖ $R_{cryst} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$, where F is the structure factor.

¶ R_{free} = R factor computed for the test set of reflections (10% for Form 1 and 3% for Form 2) omitted from the refinement.

shows that one molecule of erythropoietin binds two receptors. In both crystal forms, the arrangement of molecules is similar in the *b/c* projection, although the packing of molecules is closer and different in the *a* direction in Form 2. The most significant difference is the positioning of the D2 domains relative to each other. In Form 1 there are no receptor–receptor contacts, whereas the more closely packed Form 2 structure has a hydrogen-bonding contact between the side chains of residues Ser 135 and Glu 134 on adjacent EPObp D2 domains. This is in sharp contrast to the analogous structure with human growth hormone (hGH) bound to its receptor (hGHbp), in which the D2 domains of the two receptor molecules interact extensively with a buried surface area of over 900 Å^2 (ref. 11). The Form 1 and Form 2 structures provide independent corroboration of key features of the complex. In particular, the overall structure from both crystal forms is very similar with respect to the interfaces between erythropoietin and EPObp. The angle between the adjacent EPObp D1 domains is the same, supporting the conclusion that the interfaces in the crystal structures represent those achieved by the complex at the cell surface.

Erythropoietin has an up–up–down–down four-helical bundle topology[12] with interhelical angles similar to those of the long-chain class, for example hGH and granulocyte colony-stimulating factor. However, it also contains two small antiparallel β-strands typical of the short-chain class[13], for example macrophage colony-stimulating factor, stem-cell factor, interleukin-4 and interleukin-5. One pair of antiparallel long helices, αA (residues 8–26) and αD (residues 138–161), is held together by a disulphide bridge, Cys 7 to Cys 161. The other pair, αB (residues 55–83) and αC (residues 90–112), is linked by a short loop (Fig. 1). The αD helix is slightly irregular because of a small kink at Gly 151. The short segments of amino acids from the long AB and CD crossover loops interact with each other to form an antiparallel β-sheet: β1 (residues 39–41) and β2 (residues 133–135). Several aromatic and hydrophobic residues, such as Phe 138, Phe 142, Tyr 145, Phe 148, Leu 153 and Tyr 156, on the interior face of the D-helix, pack against the non-polar side chains from the A, B and C helices to form the hydrophobic core of erythropoietin. A second disulphide bond, Cys 29 to Cys 33, links the end of the αA helix with part of the AB loop. Erythropoietin has two additional short helices, the αB′ helix (residues 47–52) orthogonal to αB and the mini-helix αC′ (residues 114–121) following αC with a 90° tilt beginning at Gly 113.

The structure of EPObp is composed of a short amino-terminal helix (~15 residues) followed by two β-sandwich domains, D1 (N-terminal) and D2 (carboxy-terminal), each consisting of ~100 residues. The overall fold of EPObp is as described previously for the crystal structure of the erythropoietin receptor (EBP) complexed to a peptide agonist (EMP1)[6]. The N-terminal helix (residues 9–22) of both EPObp structures is tucked into the elbow formed by the D1 and D2 domains (Figs 1 and 2a), in close proximity to the WSXWS motif (residues 209–213; X = Glu) in D2. The side chain of helix residue Leu 18 makes hydrophobic contacts (<4.5 Å) with Phe 29, Leu 27, Leu 120 and Phe 208 of EPObp. Alteration of the WSXWS sequence disrupts erythropoietin binding and receptor signalling[14], correlating with similar findings for the interleukin-2, prolactin and hGH receptors[15]. The WSXWS motif has been shown to be critical for the folding and transport of the receptor to the cell surface, and an A211E mutation further improved the efficiency of the processes[8]. In the erythropoietin–(EPObp)$_2$ complex, the side chain of Glu 211 from the WSXWS motif is closest (<4.5 Å) to Leu 17 of the N-terminal receptor helix. Additionally, the side chains of residues Trp 209 and Trp 212 sandwich the hydrophobic side chain of Arg 197 in the receptor fold, whereas Ser 210 and Ser 213 are within hydrogen-bonding distance of Ala 198 and Val 196, respectively. On the basis of the interactions of the WSXWS box with the N-terminal helix and the β-sheet residues Val 196, Arg 197 and Ala 198 in the erythropoietin–(EPObp)$_2$ complex, we suggest that the N-terminal helix is biologically important in stabilizing the folded EPOR. This helix position differs from the N-terminal helices seen in the two-fold symmetric EMP1–EBP complex, in which the connection between helix and D1 was not visible and unconnected helices were farther away from the D1/D2 elbow (Fig. 2b)[6].

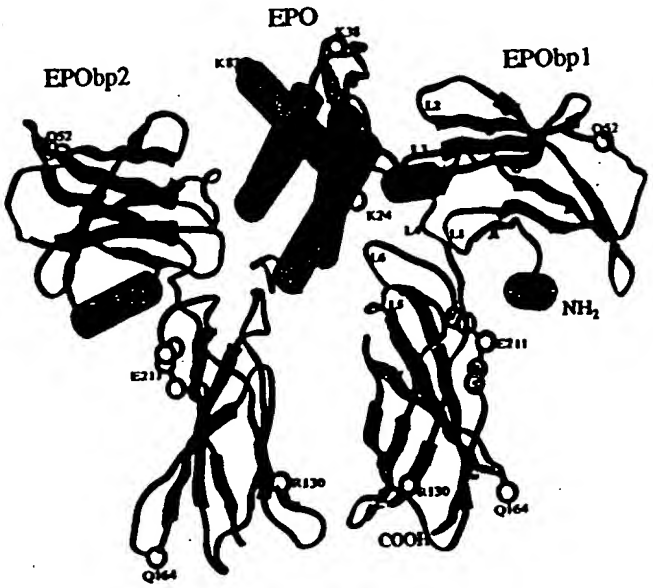


Figure 1 Crystal structure of the erythropoietin–(EPObp)$_2$ complex. α-Helices are shown as red cylinders and β-sheets as green ribbons. Erythropoietin: αA(8–26) up, αB′(47–52), αB(55–83) up, αC(90–112) down, αC′(114–121), αD(138–161) down, β1(39–41), β2(133–135). Disulphide bridges: Cys 7–Cys 161, Cys 29–Cys 33. D1 domain of EPObp: α1(9–22), βA(26–30), βB(36–42), βC(52–59), βC′(65–68), βD(69–74), βE(78–85), βF(95–103), βG(106–113). Disulphide bridges: Cys 28–Cys 38, Cys 67–Cys 83. D2 domains of EPObp: βA(119–132), βB(137–143), βC(154–162), βC′(170–174), βE(179–183), βF(190–201), βG(216–220). EPObp loops that interact with erythropoietin are labelled L1(31–35), L2(60–64), L3(98–94), L4(114–118), L5(144–153) and L6(202–205). The locations of residues in the WSXWS box, dimerizing mutation site[20] R130C, two Asn → Gln mutation sites (52, 164) on EPObp, and three Asn → Lys mutations (24, 38, 83) on erythropoietin are depicted as white spheres. EPO, erythropoietin.

In the erythropoietin–(EPObp)$_2$ complex, the receptor molecules are held together through two regions located on opposite faces of erythropoietin (Figs 1 and 2a and Table 2). These erythropoietin–EPObp interfaces have been identified as high-affinity ($K_d \approx 1$ nM) and low affinity ($K_d \approx 1$ μM) and are referred to as site 1 and site 2 respectively[4]. Site 1 is composed of erythropoietin residues from helices αA, αB′, αD and part of the AB loop. Site 2 erythropoietin residues are located in αA and αC. Complementary receptor interfaces involve contributions from six EPObp loops (L1–L6) between β-strands from both domains: L1, L2 and L3 (D1 domain); L4 from the polypeptide D1–D2 linker; and L5 and L6 (D2 domain). Residues from all six loops are involved in the site 1 interface and all except the fourth loop participate at site 2.

Site 1 is characterized by a central hydrophobic binding pocket, flanked at opposite ends by hydrophilic interactions (Table 2). EPObp loops L1, L5 and L6 accommodate the hydrophobic erythropoietin residues, with Phe 93 of EPObp dominating the non-polar interactions. This pivotal residue is firmly held in place by hydrogen bonding to erythropoietin residues Thr 44 and Asn 147. The plane of the phenyl ring of Phe 93 packs against the Cα of Gly 150 (at a distance of 3.8 Å) and is accommodated at a hydrophobic surface created by a quartet of erythropoietin side chains (Gly 151, Val 46, Phe 48 and Leu 155). Mutagenesis has shown that Phe 93 of the receptor is a critical determinant of erythropoietin binding; the P93A mutation eliminates any detectable binding[16,17]. The site

interface also has contributions from the A–B linker polypeptide residues (Thr 44–Phe 48) of erythropoietin. A smaller contribution is made by Met 150 of EPObp, which is within van der Waals contact distance of only a single amino acid, Phe 48, of erythropoietin. The hydrophobic binding surface (site 1) itself is surrounded by hydrophilic interactions (Table 2). Asn 147 of erythropoietin is central to the hydrophilic D helix interactions. It has three hydrogen bonds to EPObp residues, one of which is between its Nδ atom and the carbonyl oxygen of Phe 93. On either side of the Asn 147 are arginine residues, Arg 143 and Arg 150, that interact with glutamic acids Glu 60 (L2) and Glu 117 (L4), respectively. Flanking the binding site at one end, Ser 9 and Arg 10 from the A helix of erythropoietin interact with Glu 176 and His 153, located at the end of βC′ and in the L5 loop of the receptor, respectively. At the opposite end, Arg 143 and Lys 140 of αD, and Lys 45 of the AB loop of erythropoietin, interact with residues Glu 60, Asp 61 and Glu 62 of the L2 loop of the EPObp D1 domain.

At site 2, the interactions are less extensive than at site 1. Most of the interactions are between residues of the C helix of erythropoietin and the L3 loop of EPObp. The hydrophobic surface is created on EPObp by residues Phe 93, Phe 205 and Met 150, although because of the positioning of the D1 domain relative to the interaction of site 2 with erythropoietin there is a relatively flat surface that erythropoietin residues can interact with. The closest non-polar contact is between erythropoietin Leu 108 and EPObp Phe 93 Cγ (3.9 Å). The rest of the C helix residues are positioned at a greater distance from the EPObp hydrophobic surface than at the site 1 interface. Phe 93 is again the major contributor to the central hydrophobic binding pocket (Leu 5, Val 11, Tyr 15, Ser 104, Thr 107 and Leu 108 of erythropoietin are within 4.5 Å of EPObp Phe 93). Met 150 of EPObp, however, is more buried than in site 1, making van der Waals contacts with Arg 10, Val 11 and Arg 14 of erythropoietin. A total of 12 hydrophilic interactions are formed between residues in the A and C helices of erythropoietin (Asp 8, Arg 14, Lys 97, Ser 100,

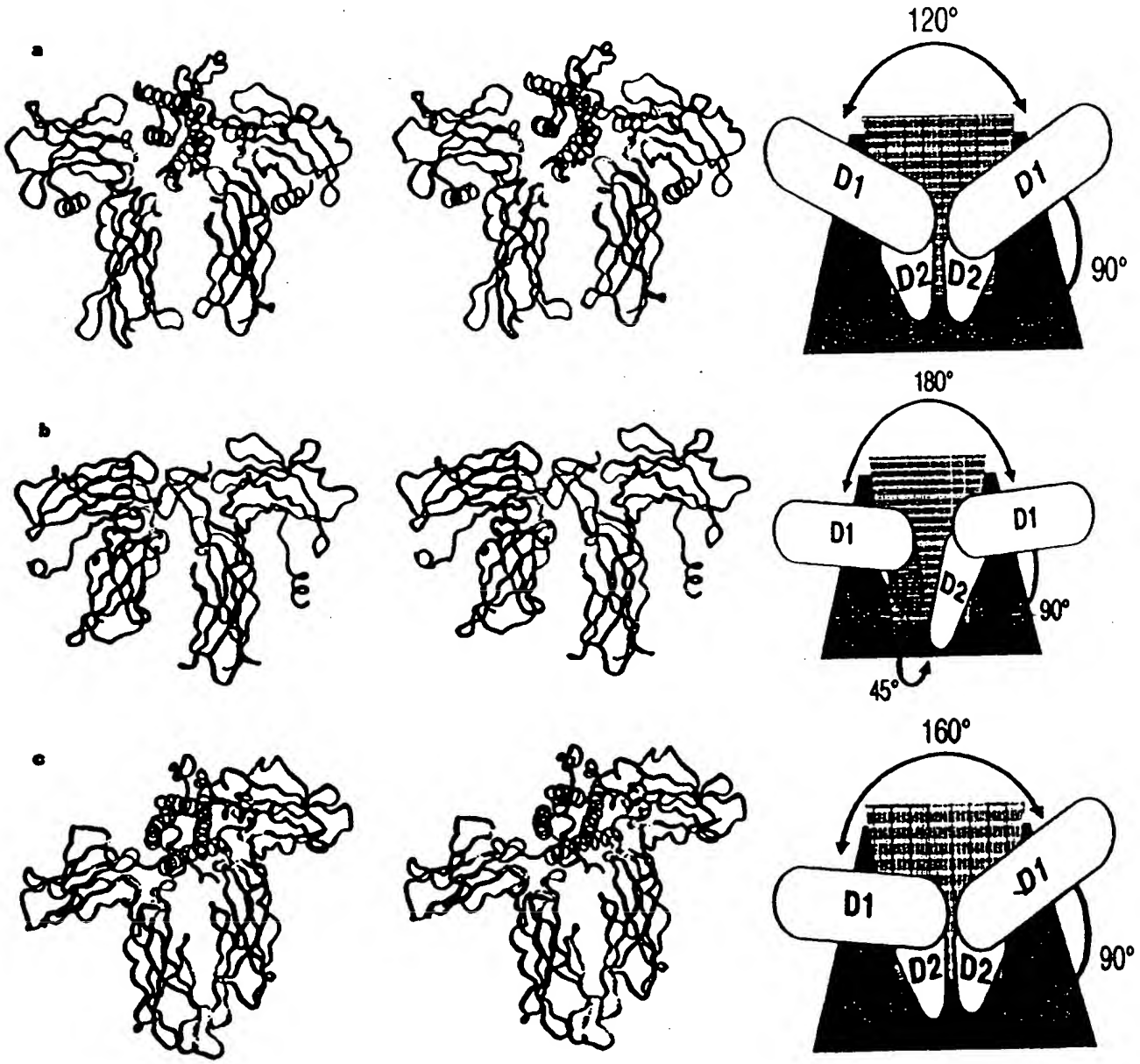


Figure 2 Stereo views. a, erythropoietin-(EPObp)$_2$. b, (EMP1-EBP)$_2$. c, hGH-(hGHbp)$_2$. All the receptors are coloured blue, with their ligands in red. The intermolecular contacts (<4.5 Å) involving receptor amino acids are shown in yellow; ligand amino acids with the receptor on the right-hand side (site 1 for erythropoietin and hGH) are shown in green; and ligand amino acids with the receptor on the left-hand side (site 2 for erythropoietin and hGH) are shown in white. Receptor orientations are also represented schematically at the right, with the horizontal membrane plane shown in grey and the corresponding vertical plane drawn in a chequered pattern.

Table 2 Hydrogen bonds and salt bridges between erythropoietin (EPO) and EPObp in sites 1 and 2 of the EPO-(EPObp)$_2$ complex

Site 1 atoms			Site 2 atoms		
EPO	EPObp	Distance (Å)	EPO	EPObp	Distance (Å)
Lys 20 Nζ	Glu 202 Oε1	3.5	Asp 8 Oδ2	His 153 Nε2	2.9
Thr 44 Oγ	Phe 93 O	2.6	Arg 14 Nη2	Leu 33 O	3.1
Lys 45 Nζ	Glu 62 Oε1	3.1	Lys 97 Nζ	Glu 34 Oε1	2.7
Val 46 N	Ser 92 O	3.0	Ser 100 O	Ser 91 Oγ	2.9
Val 46 O	Ser 92 N	2.8	Arg 103 Nε	Glu 62 Oε1	3.1
Asn 47 Nδ2	Thr 87 O	2.9	Arg 103 Nε	Glu 62 Oε2	3.5
Asn 47 Nδ2	Ala 88 O	3.2	Arg 103 Nη1	Ser 91 Oγ	2.7
Arg 131 Nη2	Asp 61 O	2.8	Arg 103 Nη1	Ala 88 O	2.6
Lys 140 Nζ	Asp 61 Oδ2	3.0	Arg 103 Nη2	Ala 88 O	3.3
Arg 143 Nη2	Glu 60 Oε2	2.8	Arg 103 Nη2	Asp 89 Oδ1	3.1
Asn 147 Nδ2	Phe 93 O	3.1	Ser 104 Oγ	Ser 92 O	2.9
Asn 147 Nδ2	His 114 Nε2	3.0			
Asn 147 Oδ1	His 114 Nε2	3.5			
Arg 150 O	Ser 204 Oγ	2.6			
Arg 150 Nη1	Pro 203 O	2.8			
Arg 150 Nη1	Glu 117 Oε2	3.2			
Arg 150 Nη2	Glu 117 Oε2	2.8			

Arg 103 and Ser 104) and loops L1, L2, L3 and L5 of EPObp (Leu 33, Glu 34, Glu 62, Ala 88, Asp 89, Ser 91, Ser 92 and His 153), with a total of 11 hydrogen bonds (Table 2). The side chains of helix A residues Asp 8 and Arg 14 form hydrogen bonds with the side chain of His 153 and the carbonyl oxygen of Leu 33, respectively.

In general, the sidechain interactions at both interfaces are predominantly between positively charged lysine and arginine residues of erythropoietin and negatively charged aspartate and glutamate side chains of EPObp. Site 1 contains almost twice as many sidechain–sidechain interactions as site 2. However, the numbers of hydrophilic contacts involving main-chain atoms from either erythropoietin or EPObp are equal for both sites. In addition, site 1 contains two hydrogen bonds that involve main-chain atoms from both erythropoietin and EPObp. Hydrophilic interactions from three EPObp residues (Glu 62, Ala 88 and Ser 92) are common to both interfaces, two of which involve hydrogen bonding from a main-chain atom (Table 2). Although most residues that interact across both interfaces are polar (Table 3), the expected contribution to binding affinity is greatest from the hydrophobic contacts.

Although the crystal structure of erythropoietin–(EPObp)$_2$ clearly shows an extensive interface between erythropoietin and its receptor (buried surface areas at sites 1 and 2 are 920 and 660 Å^2, respectively), a much smaller subset of residues, identified by mutagenesis, delineates a functional epitope (Table 3)[18–21]. Gly 151 has an important structural role in erythropoietin, forming a kink in the D helix that brings the side chain of Lys 152 into hydrophobic contact with Val 63, Trp 51 and Phe 148 in the protein core. Replacement of alanine at either position 151 or 152 causes a substantial loss of bioactivity[19–21]. The basic residues Lys 20 and Lys 45 on erythropoietin in the interface of site 1 are unaffected by Ala substitutions[19], but mutations to acidic residues do cause substantial loss of bioactivity[21]. Arg 103 of erythropoietin has been implicated in site 2 binding[18–21] and from the structure we see that it is completely buried at the interface, interacting with six residues on EPObp. Erythropoietin mutant R103A has been shown to have unchanged site 1 affinity, but to lack proliferative function owing to loss of site 2 binding[20]. A sequential binding mechanism was proposed on the basis of this analysis of several erythropoietin mutants. Arg 14 on erythropoietin, which interacts with four residues on EPObp, is also very sensitive to mutagenesis[19–21]. Sedimentation equilibrium data have shown that the R14Q mutation in erythropoietin reduces the site 2 affinity roughly fivefold (J. Philo, personal communication).

Comparison of the erythropoietin–(EPObp)$_2$ interactions with those of hGH–(hGHbp)$_2$ (ref. 11) reveals a number of significant

Table 3 List of the buried residues (>5.0 Å^2) (one-letter codes) of erythropoietin EPO in site 1 and site 2 intermolecular contact areas

Site 1				Site 2			
EPO residue	Buried surface area (Å^2)*	Neighbouring EPObp residues (4.5-Å cutoff)	Effect of EPO mutagenesis on bioactivity†	EPO residue	Buried surface area (Å^2)*	Neighbouring EPObp residues (4.5-Å cutoff)	Effect of EPO mutagenesis on bioactivity†
S9	17	H153	A, N	L5	55	F93, S204	
R10	24	H153, E176	A, I	D8	23	H153	S
E13	30	P203	A	R10	51	M150, S152, H153	A, I
L16	22	P203, S204	A, S	V11	49	F93, M150, H153, S204	S
L17	15	P203	A, S	R14	108	L33, E34, P149, M150	L, A, E, Q
K20	48	E202, P203	A, I, R, E	Y15	16	S92, F93	F, I
T44	19	S92, F93, V94	I	Q78	11	A88	A, E, R
K45	107	E62, S91, S92, F93, V94	A, R, I, D	D96	44	T87, A88	A, R
V46	38	T90, S91, S92, F93	A	K97	36	E34	R, A, E
N47	67	T87, A88, D89, T90, S91	A	V99	7	A88	A, S
F48	119	L33, E34, S92, F93, F205	I, S, G	S100	40	T87, A88, T90, S91	A, R, E, T
Y49	61	E34, T87	F, S	R103	118	L59, E62, A88, D89, S91, V94	K, A, E, H, N, Q
K52	17	E34	S, R, Q, E	S104	38	S91, S92, F93	A, I
R131	34	E60, D61, E62	T	T107	50	F93, V94	A, L, S
I133	22	D61		L108	15	F93	A, K
K140	31	E60, D61	A, R, M, T	R110	26	E60, D61	T, E
R143	43	E60, P95	E, A				
N147	43	F93, V94, P95, H114	A, K				
R150	128	F93, H114, N116, E117, P203, S204	A, Q, E				
G151	14	F93	A				
K154	32	H153, S204	A, R, S				
L155	26	F93	A, M				

*The solvent-accessible surface areas were calculated by using XPLOR3.851 with 1.80-Å probe sphere and standard atomic radii.
†Available mutagenesis data[18–20] shown with substituted amino acids highlighted according to degree of loss of *in vitro* bioactivity: bold underlined, >50 times; bold, >5 times; underlined, 2–5 times; unhighlighted, no effect.

differences. In hGH–$(hGHbp)_2$, two different regions of the AB loop of hGH (the start of the loop at Gln 46 and the end of the loop at Arg 64) are involved in site-1 ligand–receptor interactions. In contrast, only the C-terminal end of the AB loop of erythropoietin, at residue Arg 45, is involved in site-1 erythropoietin–EPObp binding. In the hGH–$(hGHbp)_2$ complex, two hGHbp residues, Trp 104 and Trp 169, form a central hydrophobic pocket flanked by hydrophilic interactions. In comparison, in the erythropoietin–$(EPObp)_2$ complex, Phe 93, the direct homologue of Trp 104 in hGHbp, is the primary contributor to the interactions between receptor and ligand in the hydrophobic pocket; Met 150 on EPObp (homologous to Trp 169 on hGHbp) is more peripherally involved. As a result of these differences, the binding between ligand and receptor mediated by the central hydrophobic pocket of the receptor seems to be more localized for erythropoietin–$(EPObp)_2$ than for hGH–$(hGHbp)_2$. This region has been identified as a mutational hot spot in hGH–$(hGHbp)_2$ (ref. 22).

Comparison of the erythropoietin–$(EPObp)_2$, $(EMP1–EBP)_2$ and hGH–$(hGHbp)_2$ structures reveals that the receptor molecules have significantly different orientations relative to each other (Fig. 2a–c). Viewed perpendicular to the membrane plane, the D1 domains from the two receptor molecules in the erythropoietin–$(EPObp)_2$ complex are positioned ~120° relative to each other, compared with 180° in $(EMP1–EBP)_2$. For the hGH–$(hGHbp)_2$ complexes, the angle is 160°. The two D2 domains in the erythropoietin–$(EPObp)_2$ and hGH–$(hGHbp)_2$ complexes are in approximately the same plane, perpendicular to the membrane. In contrast, the two D2 domains in the $(EMP1–EBP)_2$ complex are twisted away from each other by ~45°.

Disulphide crosslinking between murine erythropoietin receptors, generated by the replacement of Arg 130 with cysteine, resulted in constitutive activation[23]. The two EPObp Arg 130 residues, located near the membrane-proximal region of the D2 domains in the two adjacent receptors, are 27 Å apart. Formation of a disulphide bond would therefore require significant angular rearrangement of the receptor domains. Signalling in this case might be the result of a subset of the receptor molecules dimerizing intracellularly before being transported to the cell surface, as only a fraction of EPOR achieves constitutive activation[23]. Alternatively, a constrained angular relationship in the mutant murine dimers could bring about an impaired signalling event while still allowing erythropoietin to bind to its receptors.

The efficiency of signalling by erythropoietin mimetics is typically at most 1/20 that of the natural ligand[3]. Our results therefore show that the activating efficiency of EPOR, and by inference each of the cytokine receptor complexes, depends critically on the separation, orientation and relative disposition of bound receptors, and hence the effect on their unique intracellular domains. This consideration is therefore an important parameter in the design of erythropoietin mimetics; specifically, it suggests that asymmetrical molecules might be required to achieve these orientations most efficiently. □

Methods

Protein preparation and crystallization. Recombinant human erythropoietin mutants, r-met lys hu erythropoietin N24K N38K N83K and r-met lys hu erythropoietin N24K N38K N83K P121N P122S, were produced in *E. coli* cells and purified as described[24]. Recombinant human EPObp (r-REF huEPOR 152Q N164Q A211E) was secreted from *Pichia pastoris*, and purified as described (H.Z. *et al.*, manuscript in preparation). EPObp contains three N-terminal residues (Arg-Glu-Phe) that replace Ala in the wild-type sequence[25]. Both erythropoietin–$(EPObp)_2$ complexes were crystallized by hanging-drop vapour diffusion at 20 °C. Protein was mixed with an equal volume of the crystallization-well buffer containing 15% PEG 4K, 0.2 M $CaCl_2$ and 0.1 M MES pH 6.5–7.0 (Form 1) and in cryoprotectant conditions of 32% PEG 1500, 0.28 M $(NH_4)_2SO_4$, 0.1 M MES, pH 6.5 (Form 2). Form-1 data from three isomorphous crystals (collected at 20 °C) constituted the initial native data set. Data were collected with an R-AXIS II image plate system mounted on a Rigaku RU200 generator using a rotating copper anode target tube operating at 50 kV and 60 mA. These were autoindexed, integrated, scaled and merged with the R-AXIS associated data-reduction software BIOTEX (Molecular Structure Corporation). Form-2 data were collected with a single crystal at 100K on an R-Axis IV imaging system installed on an 18 kW Rigaku RU300 generator. Data were indexed, integrated and scaled with the programs DENZO and SCALEPACK[26].

Structure determination and refinement. For Form 1, two highly anisotropic mercury sites from the single derivative (Table 1) were refined with MLPHARE[27] and double-difference maps (mean figure of merit = 0.31). Density modification, with SOLOMON[9] and DM[27], yielded a map at 4 Å resolution that outlined the α-helices in erythropoietin and the pattern of β-sheets in the ordered chains of EPObp. 300 cycles of density modification followed as the resolution was extended from 4.0 to 2.8 Å in 0.004-Å steps. At each step the solvent mask was recalculated by using a radius equal to the resolution of the map. Structure factors calculated from the modified map were used to estimate the phases in the next resolution shell. These were not combined with the SIR phase probabilities owing to the low phasing power beyond 4 Å. Further to resolution extension, the density for domains D1 and D1′, D2 and D2′ were averaged to improve the phases iteratively by imposing constraints of equivalence in the domains. The D1 and D2 domain boundaries for each receptor were identified by inspection and the rotation angles between similar domains were determined by maximizing the correlation between the selected map regions. The initial structure solution gave an R_{free} of 46%. The structure was refined with XPLOR3.851 (ref. 28) interleaved with density averaging between the D1/D1′ and D2/D2′ domains and calculation of omit maps. Several rounds of refinement and building to 'complete' omit and $2F_o - F_c$ maps resulted in the final model (R_{cryst}/R_{free} = 19.0%/33.0%). Towards the latter stages of refinement, the release of the $(EMP1–EBP)_2$ coordinates (Brookhaven Protein Data Bank (PDB) accession number 1EBP) allowed the positioning of some surface exposed loops in the EPObp receptor.

For Form 2, heavy-atom sites were identified for each derivative in difference Patterson maps, refined for their positions and occupancies, and confirmed by examining the peaks in cross-phased difference Fourier maps[29]. The MIRAS phases calculated[29] by using isomorphous and anomalous differences from the three derivatives (Table 1) with the native data (mean figure of merit 0.72, 50.0–2.3 Å) and solvent flattened with SIGMAA weighting (mean figure of merit 0.92, 50.0–2.3 Å) produced a good-quality map. This map was used to build the complete model of erythropoietin–$(EPObp)_2$ with the program package QUANTA97 (Molecular Simulations). The starting templates for the D1 and D2 domains of EPObp used in model building were taken from the preliminary crystal structure of the $(EMP1–EBP)_2$ complex (T.D.O. *et al.*, manuscript in preparation) reproduced from previously published work[4]. The locations of cysteine residues were confirmed by examining an anomalous map, which showed clear peaks for sulphur atoms contoured at 2.0σ. The structure was refined with 705 reflections (3%) excluded for cross-validation (R_{free}) by using simulated annealing accompanied by the bulk solvent correction protocol in XPLOR3.851 (ref. 28). The model after nine cycles of manual model building, simulated annealing, and positional and temperature factor refinements yielded an R_{cryst} and an R_{free} of 0.236 and 0.339 (50.0–2.3 Å, $F > 2\sigma_f$), respectively. An additional three cycles of model building, refinement and phase combination at 1.9 Å resolution with 1323 reflections (3%) excluded for cross-validation (R_{free}) resulted in a crystallographic R_{cryst}/R_{free} of 0.242/0.318 (50.0–1.9 Å, $F > 2\sigma_f$).

Received 20 May; accepted 11 August 1998.

1. Graber, S. E. & Krantz, S. B. Erythropoietin and the control of red blood cell production. *Annu. Rev. Med.* 29, 51–66 (1978).
2. Damen, J. E. & Krystal, G. Early events in erythropoietin-induced signaling. *Exp. Hematol.* 24, 1455–1459 (1996).
3. Johnson, D. L. *et al.* Identification of a 13 amino acid peptide mimetic of erythropoietin and description of amino acids critical for the mimetic activity of EMP1. *Biochemistry* 37, 3699–3710 (1998).
4. Philo, J. S., Aoki, K. H., Arakawa, T., Narhi, L. O. & Wen, J. Dimerization of the extracellular domain of the erythropoietin (EPO) receptor by EPO: one high-affinity and one low-affinity interaction. *Biochemistry* 35, 1681–1691 (1996).
5. Wrighton, N. C. *et al.* Small peptides as potent mimetics of the protein hormone erythropoietin. *Science* 273, 458–464 (1996).
6. Livnah, O. *et al.* Functional mimicry of a protein hormone by a peptide agonist: the EPO receptor complex at 2.8 Å. *Science* 273, 464–471 (1996).
7. Derby, P., Aoki, K. H., Kuta, V. & Rohde, M. F. in *Techniques in Protein Chemistry VII* (ed. Marshak, D. R.) 109–119 (Academic, San Diego, 1996).
8. Hilton, D. J., Watowich, S. S., Katz, L. & Lodish, H. F. Saturation mutagenesis of the WSXWS motif of the erythropoietin receptor. *J. Biol. Chem.* 271, 4699–4708 (1996).

9. Abrahams, J. P. & Leslie, A. G. W. *Acta Crystallogr. D* 52, 30–42 (1996).
10. Cheetham, J. C. *et al.* NMR structure of human erythropoietin and a comparison with its receptor bound conformation. *Nature Struct. Biol.* (in the press).
11. De Vos, A. M., Ultsch, M. & Kossiakoff, A. A. Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science* 255, 306–312 (1992).
12. [illegible] & Bazan, J. F. Cytokine structural taxonomy and mechanisms of receptor engagement. *Struct. Biol.* 3, 815–827 (1993).
13. Rozwarski, D. A. *et al.* Structural comparisons among the short-chain helical cytokines. *Structure* 2, 159–173 (1994).
14. Yoshimura, A. *et al.* Mutations in the Trp-Ser-X-Trp-Ser motif of the erythropoietin receptor abolish processing, ligand binding, and activation of the receptor. *J. Biol. Chem.* 267, 11619–11625 (1992).
15. Baumgartner, J. W., Wells, C. A., Chen, C. M. & Waters, M. J. The role of the WSXWS equivalent motif in growth-hormone receptor function. *J. Biol. Chem.* 269, 29094–29101 (1994).
16. Barbone, F. P. *et al.* Mutagenesis studies of the human erythropoietin receptor. Establishment of structure–function relationships. *J. Biol. Chem.* 272, 4985–4992 (1997).
17. Middleton, S. A. *et al.* Identification of a critical ligand-binding determinant of the human erythropoietin receptor—evidence for common ligand-binding motifs in the cytokine receptor family. *J. Biol. Chem.* 271, 14045–14054 (1996).
18. Grodberg, J., Davis, K. L. & Sytkowski, A. J. Alanine scanning mutagenesis of human erythropoietin identifies four amino acids which are critical for biological activity. *Eur. J. Biochem.* 218, 597–601 (1993).
19. Wen, D. Y., Boissel, J. P., Showers, M., Ruch, B. C. & Bunn, H. F. Erythropoietin structure–function relationships–identification of functionally important domains. *J. Biol. Chem.* 269, 22839–22846 (1994).
20. Matthews, D. J., Topping, R. S., Cass, R. T. & Giebel, L. B. A sequential dimerization mechanism for erythropoietin receptor activation. *Proc. Natl Acad. Sci. USA* 93, 9471–9476 (1996).
21. Elliott, S., Lorenzini, T., Chang, D., Barzilay, J. & Delorme, E. Mapping of the active site of recombinant human erythropoietin. *Blood* 89, 493–502 (1997).
22. Clackson, T. & Wells, J. A. A hot spot of binding energy in a hormone-receptor interface. *Science* 267, 383–386 (1995).
23. Watowich, S. S. *et al.* Homodimerization and constitutive activation of the erythropoietin receptor. *Proc. Natl Acad. Sci. USA* 89, 2140–2144 (1992).
24. Narhi, L. O. *et al.* The effect of carbohydrate on the structure and stability of erythropoietin. *J. Biol. Chem.* 266, 23022–23026 (1991).
25. Johnson, D. L. *et al.* Refolding, purification, and characterization of human erythropoietin binding protein produced in *Escherichia coli*. *Prot. Express. Purif.* 7, 104–113 (1996).
26. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276, 307–326 (1997).
27. Collaborative Computational Project No. 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* 50, 760–763 (1994).
28. Brunger, A. T. *X-PLOR, Version 3.1, A System for X-ray Crystallography and NMR* (Yale Univ. Press, New Haven, 1992).
29. Furey, W. & Swaminathan, S. Phases-95: a program package for processing and analyzing diffraction data from macromolecules. *Methods Enzymol.* 277, 590–620 (1997).

Acknowledgements. We thank J. Philo, M. McGrath, P. Sprengeler, W. Welch, J. Rupert, L. Narhi, G. Rogers, M. Robde, S. Jordan, K. Langley, R. Mackman and M. Venuti for valuable discussions. Structure analyses of Form 1 and Form 2 were independently performed at Axys Pharmaceuticals Inc. and Amgen Inc., respectively.

Correspondence and requests for materials should be addressed to R.S.S. (e-mail: rsyed@amgen.com) or R.M.S. (e-mail: stroud@msg.ucsf.edu). Coordinates have been deposited with the Brookhaven Protein Data Bank (accession numbers 1hlw for Form 1 and 1eer for Form 2).

Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins

Kerstin Hill*, Kirstin Model†, Michael T. Ryan†, Klaus Dietmeier†, Falk Martin†‡, Richard Wagner* & Nikolaus Pfanner†

* *Biophysik, Universität Osnabrück, FB Biologie/Chemie, D-49034 Osnabrück, Germany*
† *Institut für Biochemie und Molekularbiologie, Universität Freiburg, Hermann-Herder-Strasse 7, D-79104 Freiburg, Germany*
‡ *Fakultät für Biologie, Universität Freiburg, D-79104 Freiburg, Germany*

The mitochondrial outer membrane contains machinery for the import of preproteins encoded by nuclear genes[1–3]. Eight different Tom (translocase of outer membrane) proteins have been identified that function as receptors and/or are related to a hypothetical general import pore. Many mitochondrial membrane channel activities have been described[4–7], including one related to Tim23 of the inner-membrane protein-import system[8]; however, the pore-forming subunit(s) of the Tom machinery have not been identified until now. Here we describe the expression and functional reconstitution of Tom40, an integral membrane protein with mainly β-sheet structure. Tom40 forms a cation-selective high-conductance channel that specifically binds to and transports mitochondrial-targeting sequences added to the *cis* side of the membrane. We conclude that Tom40 is the pore-forming subunit of the mitochondrial general import pore and that it constitutes a hydrophilic, ~22 Å wide channel for the import of preproteins.

We expressed *Saccharomyces cerevisiae* Tom40 (ref. 8) in *Escherichia coli* cells (Fig. 1a, lane 2). Tom40, representing ~25% of total *E. coli* protein, accumulated in inclusion bodies and could be solubilized with urea. As Tom40 may be slightly similar to porins at the secondary structure level[9], we modified a method for the preparation and renaturation of *Rhodopseudomonas blastica* porin from *E. coli* inclusion bodies[10]. Tom40 was isolated to high purity (>95%) (Fig. 1a, lane 3). The two protein bands of minor abundance that are visible in the purified preparation of Tom40 (Fig. 1a, lane 3) are degradation products of Tom40. When diluted into the non-ionic detergent nonanoyl-*N*-methylglucamide (Mega-9), the recombinant Tom40 migrated as a single band on blue native electrophoresis and this band was indistinguishable from that of Tom40 that had been solubilized from mitochondria by Mega-9 (Fig. 1b). Circular dichroism analysis of Tom40 in urea and after dilution into Mega-9 indicated an efficient refolding of the protein upon dilution (Fig. 1c). We inserted Tom40 into liposomes by dilution of the urea-denatured protein into a mixture of Mega-9 and azolectin, with subsequent removal of the detergent. The circular dichroism spectrum of liposome-inserted Tom40 was comparable to that of Tom40 in Mega-9 (Fig. 1c). A secondary structure calculation according to ref. 11 indicated a predominance of β-sheet structure (>60%) and little α-helical structure (<5%) in the protein.

To assess whether Tom40 was properly reconstituted in the liposomes, we determined its topological orientation in liposomes in comparison to that of Tom40 in mitochondria. Treatment of mitochondria with trypsin cleaved a small peptide (relative molecular mass (M_r) ~2.5K) from Tom40 (Fig. 1d, upper panel, lanes 2, 3). The same cleavage occurred after recombinant Tom40 inserted into liposomes was treated with trypsin (Fig. 1d, lower panel, lanes 2, 3). The remaining fragment of Tom40 was protected by the membranes, as, after lysis of mitochondria or liposomes with Triton X-100, it was digested by trypsin (Fig. 1d, lane 4). Antibodies directed against the amino-terminal or carboxy-terminal regions (12 terminal amino-acid residues in each region) of Tom40 showed that the C-terminal epitope of mitochondrial Tom40 was retained in the fragment (Fig. 1d, upper panel, lanes 6, 7), whereas the N-terminal epitope was removed by trypsin (Fig. 1d, upper panel, lane 10). The same topology of Tom40 was seen in the proteoliposomes (Fig. 1d, lower panel, lanes 6, 7, 10). A quantitative assessment of the fragmentation of Tom40 in mitochondria and liposomes (including a consideration of the higher resistance of mitochondrial membranes to trypsin) indicates that at least 50–60% of the recombinant Tom40 was inserted into liposomes with the correct orientation (another fraction of recombinant Tom40 was fully digested by trypsin even in the presence of intact liposome membranes, probably representing Tom40 associated with the liposome preparation but not inserted into the membrane). Indeed, a preprotein specifically affected the channel activity only when added to the *cis* side of the membrane (that is, to the N terminus of Tom40) (see below). We conclude that the recombinant Tom40 molecules that were functionally incorporated into liposomes were asymmetrically inserted with the same relative orientation as in the outer mitochondrial membrane.

To determine whether reconstituted Tom40 could form a channel, we performed electrophysiological studies. After fusion of Tom40-containing liposomes with a planar bilayer, single-channel currents could indeed be measured (Fig. 2a). At membrane potentials below 100 mV, the channels were mainly open and we observed brief, flickering closures of the channel. When examined at higher time resolution (10 kHz), the current recordings showed subconductance levels (lower trace, Fig. 2a). The current traces show that